

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2005 年 12 月 8 日 (08.12.2005)

PCT

(10) 国際公開番号  
WO 2005/117366 A1

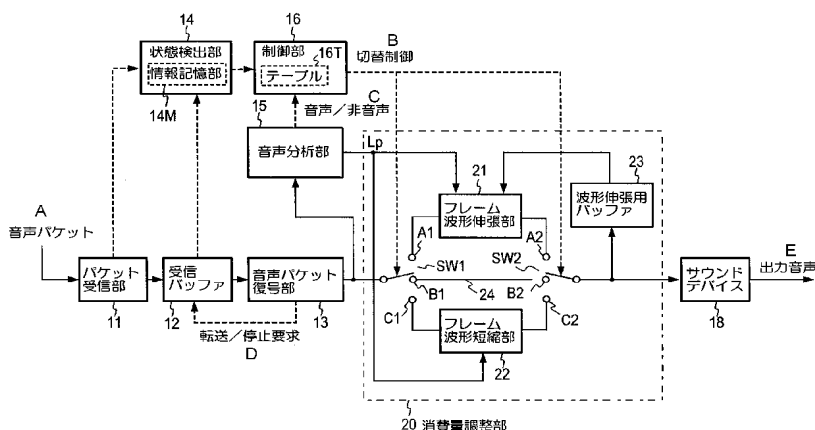
- (51) 国際特許分類<sup>7</sup>: H04L 12/56, G10L 19/00  
 (21) 国際出願番号: PCT/JP2005/009569  
 (22) 国際出願日: 2005 年 5 月 25 日 (25.05.2005)  
 (25) 国際出願の言語: 日本語  
 (26) 国際公開の言語: 日本語  
 (30) 優先権データ:  
 特願2004-156069 2004 年 5 月 26 日 (26.05.2004) JP  
 (71) 出願人 (米国を除く全ての指定国について): 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町二丁目 3 番 1 号 Tokyo (JP).

- (72) 発明者; および  
 (75) 発明者/出願人 (米国についてのみ): 大室 伸 (OHMURO, Hitoshi) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目 9 番 1 1 号 N T T 知的財産センタ内 Tokyo (JP). 森 岳至 (MORI, Takeshi) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目 9 番 1 1 号 N T T 知的財産センタ内 Tokyo (JP). 日和▲崎▼祐介 (HIWASAKI, Yusuke) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目 9 番 1 1 号 N T T 知的財産センタ内 Tokyo (JP). 片岡 章俊 (KATAOKA, Akitoshi) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目 9 番 1 1 号 N T T 知的財産センタ内 Tokyo (JP).

[続葉有]

(54) Title: SOUND PACKET REPRODUCING METHOD, SOUND PACKET REPRODUCING APPARATUS, SOUND PACKET REPRODUCING PROGRAM, AND RECORDING MEDIUM

(54) 発明の名称: 音声パケット再生方法、音声パケット再生装置、音声パケット再生プログラム、記録媒体



- 14 STATE DETERMINING PART  
 14M INFORMATION STORING PART  
 16 CONTROL PART  
 16T TABLE  
 B SWITCH CONTROL  
 C SOUND/NON-SOUND  
 15 SOUND ANALYZING PART  
 A SOUND PACKET  
 11 PACKET RECEIVING PART  
 12 RECEPTION BUFFER  
 D REQUEST TRANSFER/HALT  
 13 SOUND PACKET DECODING PART  
 21 FRAME WAVEFORM EXTENDING PART  
 22 FRAME WAVEFORM COMPACTING PART  
 23 WAVEFORM EXTENSION BUFFER  
 18 SOUND DEVICE  
 E OUTPUT SOUND  
 20 CONSUMPTION AMOUNT ADJUSTING PART

(57) Abstract: A received packet is stored in a reception buffer, while a state determining part determines the maximum delay fluctuation and the buffering amount of the reception buffer at the arrival of the packet. A control part decides, from a prescribed table, an optimum buffering amount associated with the maximum delay fluctuation, and further determines, from the determined buffering amount and the optimum buffering amount, an urgency of the buffering amount adjustment. A consumption amount adjusting part performs, based on the urgency, the determined buffering amount and the optimum buffering amount, waveform extension and compaction for the decoded sound data row of a current frame as decoded from the packet read from the reception buffer, thereby adjusting the consumption amount of the reproduced frame so as to prevent the reception buffer from being vacant.

(57) 要約: 受信パケットを受信バッファに蓄積すると共に状態検出部によりそのパケットの到着時の最大遅延ゆらぎと受信バッファのバッファ量とを検出し、制御部により最大遅延ゆらぎに対応する最適バッファ量を予め決めた表から求め、検出バッ

ファ量と最適バッファ量からバ

[続葉有]



(74) 代理人: 中尾 直樹, 外(NAKAO, Naoki et al.); 〒1600022 東京都新宿区新宿三丁目 1 番 2 2 号 新宿 NSビル 4 階 Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告書

2 文字コード及び他の略語については、定期発行される各 PCT ガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

ツファ量の調整の緊急度を判定し、受信バッファから読み出したパケットから復号された現フレームの復号音声データ列に対し、消費量調整部において上記緊急度と、検出したバッファ量と、上記最適バッファ量とに基づいて上記復号音声データ列に対し波形の伸張、短縮を行って再生フレームの消費量を調整することにより、受信バッファが空になることを阻止する。

## 明 細 書

音声パケット再生方法、音声パケット再生装置、音声パケット再生プログラム、記録媒体

## 技術分野

[0001] この発明は、デジタル化された音声、音楽などの音響信号(以下総称して音声信号)をインターネットをはじめとするパケット通信網を介して送信する際に、受信側において安定した品質で音声信号を再生するために用いる音声パケット再生方法、音声パケット再生装置、音声パケット再生プログラム及びこのプログラムを記録した記録媒体に関する。

## 背景技術

[0002] 音声信号をVoice over IP技術(以下単に音声パケット通信と称す)を利用して送信し受信するサービスが普及しつつある。図1はその概要を示す。入力音声信号は音声信号送信装置5で音声パケットに変換され、パケット通信網6に送出される。音声信号受信装置7は自己宛に送られて来る音声パケットを識別して受信し、音声パケットを復号して音声を出力する。

図2に送信すべき音声データ列と音声パケットの関係を示す。図2Aは送信すべき音声データ列を示す。送信すべき音声データ列は一般にPCMデジタルサンプル列で構成される。このデジタル音声データ列をフレームと呼ばれる一定の時間単位(一般的には10ミリ秒～20ミリ秒程度)に区切って音声符号化し、音声符号に送信時刻を表わすタイムスタンプ等を付加して音声パケットとして送信される。音声パケットは図2Bに示すように時間軸方向に圧縮された間欠信号とされ、間欠信号の空き時間は他のパケット通信に利用される。音声信号送信装置5から送り出される音声パケットの送り出しの時間の間隔は音声データ列のフレーム長に該当し、フレーム長に相当する時間間隔でパケット通信網6に送り出される。

[0003] 音声信号受信装置7はフレーム長に相当する時間間隔で到来する音声パケットを受信し、音声パケットを復号することにより図2Cに示すように1音声パケットが1フレーム長の音声データ列に復号され、1フレーム長に相当する時間間隔で音声パケットを

受信することにより、音声信号受信装置7は連続した音声を再生することができる。

ところで、通信網の状態によってはパケットの到来時間に大きなゆらぎが生じ、その結果、制限時間内(フレーム長に相当する時間内)にパケットが届かない現象が発生し、再生する音声が続切れという問題がある。この問題に対して、ゆらぎ吸収バッファとも呼ばれる受信バッファを設けて、常に一定量のパケットを受信バッファに溜めておく方法が知られている。このとき、受信バッファに溜めておくパケットの量を多くすると、パケットの到着遅延の大きなゆらぎには強いが、パケットを受信してから音声再生されるまでの遅延、即ち通話の遅延も大きくなり、双方向で話しづらいという問題がある。逆に、受信バッファに溜めておくパケットの量を少なくすると、通話の遅延は少ないが、パケットの到着遅延のゆらぎが発生したときに、音声が続切れやすいという問題がある。つまり、通話の遅延と音声の途切れやすさはトレードオフの関係にある。

[0004] この問題に対して、受信バッファに溜めるパケット量をダイナミックにコントロールする方法が知られている。これは、通話開始時は受信バッファに溜めるパケット量を少なくして、通話遅延を小さくするが、通話中にバッファに溜めたパケットが枯渇すると、通話途中で一旦音声再生処理を中止して受信バッファに溜めるパケット量を一定量増加させ、以降において通話音声が続切れにくくするものである。

人間が発声する音声を10ミリ秒～20ミリ秒の時間単位に区切って見ると、通常が発声では発生時間の数十パーセント程度は非音声区間(背景雑音区間及び無音区間)であると言われている。そこで、受信バッファ中の受信パケット数が第1の閾値を超えた場合、復号音声中の非音声区間を削除することによりフレーム長を短縮して受信バッファ中の音声再生のための次のパケットへのアクセス時点を早め、受信バッファ中の受信パケット数が第1の閾値より小さい第2の閾値より小さくなった場合は、復号音声信号中の非音声区間を伸張し、それによって受信バッファ中の音声再生のための次のパケットへのアクセス時点を遅延させることにより遅延ゆらぎに対処することができる。この方法は、非音声区間の発生頻度が著しく小さくなったり、長時間にわたって非音声区間が無かったりすると、受信バッファによる調整ができなくなってしまう。

[0005] 一方、音声区間(有声音区間及び無声音区間)において、ピッチ波形を単位として

挿入、削除すれば、聴覚的な品質劣化をほとんど生じさせることなく、時間長の伸縮が可能であることが非特許文献1に記載されている。この方法を用いて前記非音声区間だけでは十分な受信バッファの調整ができない問題を解決するため、受信バッファ内の蓄積パケット量が下限値より下がると音声区間においてピッチ長の補間音声波形を追加し、上限値を超えると音声区間においてピッチ長の音声波形を間引くことが特許文献1に示されている。しかしながら、ピッチ長の波形を挿入あるいは間引くことにより音質劣化を低減しているものの、このピッチ長の波形の挿入又は間引きは、バッファ内の蓄積パケット量が上限閾値と下限閾値の間になるまで一連のフレームに対し実施されるため、再生音声に望ましくない程度の音質劣化を与える可能性がある。また、下限値及び上限値は固定されているため、遅延ゆらぎの急激な変化に対応できず、パケットロスが生じてしまう場合もある。

特許文献1: 日本国特許出願公開2003-050598号公報

非特許文献1: 森田、板倉、“ポインター移動量制御による重複加算法 (PICOLA) を用いた音声の時間軸での伸張圧縮とその評価”、日本音響学会講演論文集、1-4-14. 1986年10月

## 発明の開示

### 発明が解決しようとする課題

- [0006] 本発明の目的は、ピッチ波形を単位とした挿入、削除処理を応用して、より高性能な音声パケット再生方法、及び装置を提案しようとするものである。

### 課題を解決するための手段

- [0007] この発明によれば、音声データ列のフレーム毎の符号化により生成された音声符号を收容して送信された一連の音声パケットを受信して再生する音声パケットの再生方法は、以下のステップを含む:
- (a) 受信パケットを受信バッファに蓄積し、
  - (b) 上記受信パケットの到着時間のゆらぎを一定期間観測して得られるゆらぎの最大値又は統計値、以下最大遅延ゆらぎと呼ぶ、と上記受信バッファに蓄積されているパケット数であるバッファ蓄積パケット数とを検出し、
  - (c) 上記最大遅延ゆらぎから、予め決めた最大遅延ゆらぎと最適蓄積パケット数の

関係を用いて、上記受信バッファに蓄積すべき最適なパケット数、以下最適蓄積パケット数と呼ぶ、を求め、

(d) 上記検出したバッファ蓄積パケット数と上記最適蓄積パケット数の相違の程度を複数の段階で判定し、

(e) 上記受信バッファから現フレームに対応するパケットを取り出し、そのパケット中の音声符号を復号して現フレームの復号音声データ列を得て、

(f) 上記相違の程度の段階ごとに定めた規則に基づいて、上記バッファ蓄積パケット数を上記最適蓄積パケット数に近づけるように、上記復号音声データ列に対しその波形の伸張、短縮又は現状維持のいずれかの処理を行い、現フレームの再生音声データとして出力する。

[0008] この発明によれば、音声データ列のフレーム毎の符号化により生成された音声符号を収容して送信された一連の音声パケットを受信して再生する音声パケットの再生装置は、

入力音声データ列のフレーム毎の符号化により生成された音声符号を収容して送信された一連の音声パケットを受信して再生する音声パケットの再生装置であり、

音声パケットをパケット通信網から受信するパケット受信部と、

受信したパケットを一時的に蓄積し、要求に応じてパケットを読み出す受信バッファと、

上記受信パケットの到着時間のゆらぎを一定期間観測して得られるゆらぎの最大値又は統計値、以下最大遅延ゆらぎと呼ぶ、と上記受信バッファに蓄積されているパケット数であるバッファ蓄積パケット数を検出する状態検出部と、

上記最大遅延ゆらぎから、予め決めた最大遅延揺らぎと最適蓄積パケット数の関係を用いて、上記受信バッファに蓄積すべき最適なパケット数、以下最適蓄積パケット数と呼ぶ、を求め、上記検出したバッファ蓄積パケット数と上記最適パケット数の相違の程度を複数の段階で判定し、上記相違の程度の段階ごとに予め定めた規則に基づいて、上記バッファ蓄積パケット数を上記最適蓄積パケット数に近づけるように波形の伸張、短縮、又は現状維持を指示する制御信号を生成する制御部と、

上記受信バッファから取り出された現フレームに対応するパケット中の音声符号を

復号して現フレームの復号音声データ列を得る音声パケット復号部と、

上記制御信号に従って、上記現フレームの復号音声データ列に対し、その波形の伸張、短縮又は現状維持のいずれかの処理を行い、現フレームの再生音声データとして出力する消費量調整部、

とを含むように構成される。

### 発明の効果

- [0009] パケット到着時間の遅延ゆらぎの大きいパケット通信網によってリアルタイムに音声信号を通信する場合に、本発明を適用することによって、音声の有無を問わずに定常的に音声データ列の消費量を制御し、受信バッファ内のパケット量を調整するから通信網の状態(ゆらぎ時間)の変化に追隨して、受信バッファの最適コントロールが可能となる。この結果、音声の途切れがなく、かつ通話遅延を必要最小限に抑えた音声通話が実現される。また、パケット通信網はコストを抑えるために、ある程度のゆらぎを許容するように設計するのが一般的であり、本発明の利用によって、回線自体のゆらぎが少ない高品質ネットワークを利用しなくても音声の途切れが発生しないため、ネットワークの利用に関するコスト削減にも効果がある。

### 図面の簡単な説明

- [0010] [図1]パケット通信の概要を説明するためのブロック図。
- [図2]パケット通信の概要を説明するためのタイミングチャート。
- [図3]本発明の音声パケット再生装置の一実施例を説明するためのブロック図。
- [図4]最大遅延ゆらぎに対する最適バッファ量を示す表。
- [図5]図5Aは受信パケットの処理手順を示し、図5Bはパケットの読み出し、復号、波形処理の手順を示す。
- [図6]図3における音声分析部15の構成例を示すブロック図。
- [図7]図3におけるフレーム波形伸張部21の機能構成例を示すブロック図。
- [図8]図7に示したフレーム波形伸張部21による波形伸張処理を模式的に説明するための波形図。
- [図9]図7に示したフレーム波形伸張部21による波形伸張処理の他の例を説明するための波形図。

[図10]図3におけるフレーム波形伸張部21の他の構成を説明するためのブロック図。

[図11]図3に示したフレーム波形短縮部22の内部を構成を説明するためのブロック図。

[図12]図11に示したフレーム波形短縮部22の動作を説明するための波形図。

[図13]図3に示した消費量調整部20の動作を説明するためのタイミングチャート。

[図14]バッファ調整の緊急度を判定する手順を示すフローチャート。

[図15]緊急度に対する消費量調整部20の制御例を示す表1。

[図16]緊急度に対する消費量調整部20の制御例を示す表2。

[図17]緊急度に対する消費量調整部20の制御例を示す表3。

[図18]緊急度に対する消費量調整部20の制御例を示す表4。

[図19]図3における音声分析部15の第2の実施例を示すブロック図。

[図20]図3における音声分析部15の第3の実施例を示すブロック図。

[図21]図20における音声フレームの区間の判定手順を示すフローチャート。

[図22]図21による判定結果の音声フレームに対する消費量調整部20の制御例を示す表5。

## 発明を実施するための最良の形態

[0011] 本発明は、コンピュータ本体とコンピュータプログラムによって実行することが可能であるし、デジタルシグナルプロセッサや専用LSIに実装して実現することも可能である。特に切替スイッチはプログラムの条件分岐として実装することができる。

### 第1実施例

図3は、本発明における音声信号受信部の構成例であり、パケット受信部11と、受信バッファ12と、音声パケット復号部13と、状態検出部14と、音声分析部15と、制御部16と、消費量調整部20と、サウンドデバイス18とから構成されている。パケット受信部11は、通信網から音声パケットを受信し、その受信バッファ12に蓄積すると共に、そのパケットの到着時刻とそのパケットのタイムスタンプを状態検出部14に与える。音声パケットには、フレームと呼ばれる一定の時間単位(一般的には10ミリ秒～20ミリ秒程度)に区切った音声信号を音声符号化の手法によって変換した音声符号と、送信側におけるパケット作成の時間順序を示すタイムスタンプ(フレーム番号に対応



が含まれている。代表的な音声符号化手段としては、ITU-T(国際電気通信連合)の標準であるG.711方式があるが、その他任意の方式でもよい。

- [0012] 受信バッファ12は、受信音声パケットを蓄積し、音声パケット復号部13から転送要求がある毎に、音声パケットを音声パケット復号部13にタイムスタンプ順に送る。受信バッファ12から音声パケットが音声パケット復号部13に転送されるごとに、受信バッファ12内のその音声パケットは破棄される。

音声パケット復号部13は、音声パケットに含まれる音声符号を復号して、音声データ列を出力し、消費量調整部20と音声分析部15に与える。ここでいう音声データ列とはデジタル音声サンプル列であり、一般的にはPCMフォーマットで扱われることが多い。以下、特に音声信号がアナログかデジタルかの表記はしないが、最終的に人間の耳で聞く音がアナログ信号である以外の処理途中の信号はデジタル信号である。

- [0013] また、送信側ではPCM信号をフレームごとに符号化してからパケットに収容して送信する場合と、PCM信号をそのままフレームごとにパケットに収容して送信する場合がある。前者の場合、受信側において受信パケットから取り出した音声符号を復号する手段を必要とするが、後者の場合は必要としない。しかしながら、PCM信号も一種の符号であると考えれば、受信側においてパケットから取り出したPCM信号をデジタルオーディオ信号に変換する復号化手段を必要とする。図3に示したこの発明の音声パケット再生装置における音声パケット復号部13はいずれの場合であってもよい。

- [0014] 状態検出部14はパケット受信部11で受信されたパケットの到着時刻とタイムスタンプが与えられ、遅延ゆらぎを検出すると共に、受信バッファ12に蓄積されているパケット数を検出して制御部16に与える。ここで簡単のため現パケットのタイムスタンプをフレーム番号 $F_n$  ( $n=0, 1, 2, \dots$ )と同じとし、到着時刻を $T_n$ とし、直前のパケットのタイムスタンプを $F_{n-m}$ とし、その到着時刻を $T_{n-m}$ とする。信号伝送路の遅延量の変動が小さければ通常は $m=1$ であるが、遅延量が大きく変動すると到着パケットの順番の逆転が生じる可能性があるので一般には $m=1$ とは限らない。フレーム長を $L_f$ とし、遅延ゆらぎ $j_n$ を例えば $j_n = (T_n - T_{n-m}) - m \times L_f$ 又はその絶対値で表すことにする。遅延ゆらぎ $j_n$ の定義

はゆらぎの程度を数値で表すことができれば他の定義式を用いてもよい。

- [0015] 状態検出部14は前回のパケットの到着時刻 $T_{n-m}$ とタイムスタンプ $F_{n-m}$ を内部の情報記憶部14Mに保持すると共に、一定期間(例えば2秒間)過去に遡って受信された各パケットについて計算した遅延ゆらぎ $j_n, j_{n-1}, \dots$ を保持している。現パケット(タイムスタンプ $F_n$ )の受信時点における過去一定期間(2秒)内の遅延ゆらぎの最大値又は統計値(以下最大遅延ゆらぎと呼ぶ)を $J$ と表すことにする。あるいは一定期間の代わりに、一定数(例えば100)の受信されたパケットに対する最大遅延ゆらぎを $J$ としてもよい。状態検出部14は、パケット受信部11からパケットの到着時刻 $T_n$ とタイムスタンプ $F_n$ が与えられる毎に最大遅延ゆらぎ $J$ を求め、現時点の受信バッファ12の蓄積パケット数(バッファ量) $S$ と共に制御部16に与える。
- [0016] 音声分析部15は、まず音声パケット復号部13から出力された復号音声データ列を分析して、当該フレームが音声区間であるか非音声区間であるかを判定する。なお、ここで言う音声区間とは、人間の発声した音声信号が含まれる区間、非音声区間とは音声信号が含まれない区間と定義する。音声区間は有声音区間又は無声音区間のいずれかであり、非音声区間は音声区間以外の区間、即ち、背景雑音区間又は無音区間のいずれかである。その判定方法としては、例えば、フレーム内の音声信号のパワーを計算し、パワーが閾値以上であれば音声フレーム(音声区間)、閾値未満であれば非音声フレーム(非音声区間)と決めることにする。
- [0017] 当該フレームが音声区間と判定された場合には、ピッチ長の分析が行なわれる。ピッチ長は、例えば、音声波形又は音声波形にスペクトル包絡の逆特性を持つフィルタをかけた信号の自己相関係数を計算することによって得られる。当該フレームが非音声区間と判定された場合には、ピッチ長の分析は行なわず、ピッチ長 $L_p$ として一定値、例えばフレーム長 $L_f$ の1/2を設定する。後述するように、ピッチ長 $L_p$ はフレーム長 $L_f$ の1/2以下が以後の処理に都合がよいためである。フレーム長 $L_f$ の1/4、1/6といった1/2以下の任意の値でもよい。
- [0018] 当該フレームが音声区間と判定された場合であっても、その音声の有声音の場合と無声音の場合がある。無声音の場合にはピッチという音声の物理的特徴量は存在しない。その場合でも、ピッチ分析の手法を用いて得られた値をピッチ長として以後

の処理に利用して差し支えないし、非音声区間と同様にピッチ長として一定値に設定してしまっても本発明の効果に大差はない。

背景雑音がある場合の背景雑音区間(音声区間以外)では、音声区間と同様にピッチ分析の手法を用いてピッチ長を算出する。このとき得られるピッチ長は、音声の物理的特徴量としてのピッチとは異なるが、信号に含まれる主要な基本周波数に対応する周期として利用する。なお、音声の符号化方法によっては音声符号中にピッチに関する情報が含まれている場合があり、その場合には音声符号中のピッチ情報を用いてピッチ長を求めればよい。

[0019] 制御部16は予め実験により最大遅延ゆらぎとそれに対する受信バッファ12の最適蓄積パケット数を決めた例えば図4に示す表16Tを有している。図4に示す表16Tは、フレーム長Lfが20msecの場合であり、従って、隣接パケットとの理想的な受信間隔も20msecである。制御部16は、状態検出部14から与えられた最大遅延ゆらぎJから表16Tを参照して求めた最適パケット数Bと現在の受信バッファ12の蓄積パケット数Sとの差 $S - B$ に基づいて、受信バッファ12に必要とされるバッファ量の調整の緊急度を判定する。なお、表16Tを用いる代わりに、例えば最大遅延ゆらぎJと最適蓄積パケット数Bの関係を近似的に表す式を予め決め、それを使用してもよい。即ち、最大遅延ゆらぎJと最適蓄積パケット数Bの予め決めた関係を使用するのであればどのような方法でもよい。

[0020] 制御部16は判定した緊急度と、音声分析部15から与えられた音声／非音声判定結果に基づいて、音声パケット復号部13からの現フレームの復号音声波形データに対し、伸張するか、短縮するか、あるいは変更しないかを決め、その決定による制御を消費量調整部20に与える。

消費量調整部20は制御部16の制御に従って、音声パケット復号部13からの復号音声波形データをそのまま出力するか、伸張して出力するか、短縮して出力する。復号音声波形を伸張すれば音声パケット復号部13の受信バッファ12に対し次のパケットの転送要求が遅れることになり、単位時間当たりのパケット消費量が少なくなる。逆に、復号音声波形を短縮すれば、単位時間当たりのパケット消費量が多くなる。即ち音声信号として出力するための単位時間当たりの処理フレーム数を制御すること

により受信バッファ12からのパケットの読み出し量を制御する。

- [0021] サウンドデバイス18は図示していないD/A変換器を有しており、音声データ列をアナログ信号に変換し実際にスピーカで再生する。消費量調整部20からデジタルの音声信号がサウンドデバイス18に送られると、アナログ音響信号である出力音声再生される。サウンドデバイス18は受け取った時間長(1フレームに相当する時間)の信号の再生が終わると、次のパケットの復号音声データ列を受け取る。サウンドデバイス18の中にもサウンドデバイス用のバッファを持つのが一般的で、ダブルバッファと呼ばれる方法がよく知られている。ダブルバッファとは、2つのバッファを持ち、一方が再生中は他方は次の再生のための信号を受け取って準備する方法である。バッファが満杯のときは、一方の再生が終了するまでは次の信号は受け取らない。バッファ内に次の信号を格納できる空き領域が発生すると、直ちに次の信号の読み込みが実行される。

- [0022] 図5A及び図5Bは図3に示したこの発明による音声パケット再生装置におけるパケット受信部11と受信バッファ12と、状態検出部14による上述した受信パケットの処理手順と、それ以外の構成部による音声データ列の再生処理手順を示す。

図5Aの処理では、ステップS1Aでパケット受信部11により音声パケットが受信されると、ステップS2Aで受信パケットのタイムスタンプと到着時刻とを状態検出部14に保存する。

ステップS3Aで、状態検出部14により受信パケットの到着時刻とタイムスタンプから直前の受信パケットに対する遅延ゆらぎを求め、到着時刻、タイムスタンプと共に保存し、過去一定時間内の最大遅延ゆらぎJを求め、受信バッファ12の現時点の蓄積パケット数(バッファ量)Sを検出し、制御部16に最大遅延ゆらぎJとバッファ量Sを与える。

- [0023] ステップS4Aで、受信パケットを受信バッファ12に蓄積し、ステップS1Aに戻って次のパケットを待つ。

図5Bの処理では、ステップS1Bで音声パケット復号部13からの要求により受信バッファ12から再生処理における現フレームに対応するパケットを読み出し、ステップS2Bで音声パケット復号部13によりパケット中の音声符号を復号化して音声データ列

を得る。

ステップS3Bで、音声分析部15により復号音声データ列が音声区間であるか非音声区間であるか判別し、ステップS4Bで制御部16により最大遅延ゆらぎJに対する最適バッファ量Bを図4の表から決定する。

- [0024] ステップS5Bで、制御部16によりさらに最適バッファ量Bと検出バッファ量Sに基づいてバッファ量の調整の緊急度を判定する。

ステップS6Bで、判定された緊急度に応じて消費量調整部20により現フレームの復号音声データ列に対し波形伸張／短縮処理を行う。

ステップS7Bで、波形伸張／短縮処理された音声データ列を出力し、ステップS1Bに戻って次のパケットの再生処理に移る。

以下、図3に示したこの発明による音声パケット再生装置の要部を詳細に説明する。

- [0025] 図6は音声分析部15の機能構成例を示す。ここでは、音声分析部15は、高域通過フィルタ151と、分析用バッファ152と、フレーム内パワー計算部153と、区間判定部154と、ピッチ抽出部15Aとから構成されている。音声パケット復号部13からのフレームごとの復号音声データは高域通過フィルタ151により直流分や、本来音声に含まれていない極端に低い周波数(例えば50Hz以下)の成分が除去された後、以降の処理のために分析用バッファ152に一時保持され、フレーム内パワー計算部153によりフレーム内のパワーP<sub>f</sub>が計算される。区間判定部154は予め決めたパワー閾値P<sub>th</sub>よりもフレーム内パワーP<sub>f</sub>が大きければ音声区間、P<sub>th</sub>以下であれば非音声区間と判定し、判定結果を出力する。

- [0026] ピッチ抽出部15Aは区間判定部154の判定結果が音声区間と判定した場合に復号音声データ列からピッチ抽出処理を実行する。ピッチ抽出部15Aは線形予測分析部155と、逆フィルタ係数算出部156と、逆フィルタ157と、ピッチ相関計算部158とから構成されている。線形予測分析部155は分析用バッファ152に保持されている1フレーム分の復号音声データ列を線形予測分析し、線形予測係数を得て逆フィルタ係数算出部156に与える。逆フィルタ係数算出部156は線形予測係数から、復号音声信号のスペクトル包絡を平坦化するような逆フィルタ係数を算出し、線形フィルタで

構成された逆フィルタ157の係数として設定する。従って、逆フィルタ157は与えられた復号音声データ列を逆フィルタ処理し、それによってスペクトル包絡が平坦化された音声データ列がピッチ相関計算部158に与えられる。ピッチ相関計算部158は与えられた音声データの自己相関値をサンプル点を順次ずらしてそれぞれ計算し、得られた一連の相関値のピークの間隔をピッチ長 $L_p$ として検出し、消費量調整部20に与える。

[0027] 尚、ピッチ長の分析には、当該フレームの信号だけでなく、過去のフレームの信号も利用することが多い。その場合には分析用バッファ152のサイズを2フレーム以上とし、現フレームと過去のフレームの復号音声データ列を保持し、過去から現在に渡る複数フレームの音声データ列に対するピッチ長分析を行えばよい。音声／非音声の判定結果は制御部16に、ピッチ長 $L_p$ は音声データ列の消費量を調整する消費量調整部20に送られる。

図3に戻って、消費量調整部20は、フレーム波形伸張部21と、フレーム波形短縮部22と、波形伸張用バッファ23と、スループス24と、スイッチSW1、SW2とによって構成される。スイッチSW1は3つの切替端子A1、B1、C1を持ち、スイッチSW2は3つの切替端子A2、B2、B3を持ち、これら切替端子の組A1、A2;B1、B2;C1、C2間にそれぞれフレーム波形伸張部21、スループス24、フレーム波形短縮部22が接続されており、制御部16の制御に従ってこれらのいずれかを選択するように連動して切り替えられる。スイッチSW2により選択された音声波形データ列はサウンドデバイス18に与えられると共に、波形伸張用バッファ23に書き込まれる。波形伸張用バッファ23から読み出された音声信号はフレーム波形伸張部21に与えられる。

[0028] フレーム波形伸張部21の機能構成例を図7に、フレーム波形伸張部21による波形伸張処理の例を図8に示す。

フレーム波形伸張部21は、波形処理用バッファ21-0と、波形挿入処理部21-1と、第1波形切出し部21-2と、ピッチ波形生成部21-3と、第2波形切出し部21-4とから構成されている。第2波形切出し部21-4は音声分析部15からのピッチ長 $L_p$ を使って、波形伸張用バッファ23に蓄積されている1フレーム過去の出力音声信号波形WF1から、図8、行Aに示すピッチ長 $L_p$ の区間の波形Xを切り出す。即ち、バッファ23内の最

後のサンプル点から、過去の時間方向に向かって1ピッチ長 $L_p$ の波形Xを切り出す。

- [0029] 第1波形切出し部21-2は、波形処理用バッファ21-0内の現フレームの音声信号波形WF0から、図8の行Aにおけるピッチ長 $L_p$ の区間の波形Yを切り出す。即ち、入力された現フレームの音声信号波形WF0の先頭サンプル点から正の時間方向に向かって1ピッチ長 $L_p$ の波形を切り出す。

ピッチ波形生成部21-3は、前記切り出した波形Xと波形Yにそれぞれ三角窓の重み付けを行なった後、互いに加算して、図8の行Bに示す波形Zを作成する。ここで利用する重み付け三角窓は、非特許文献1に記載されている三角窓と同様のものを利用することができる。即ち、波形区間の始点から終点に向かって、重みが波形Xでは0から1に、波形Yでは1から0に直線的に変化する形状を利用できる。

- [0030] 波形挿入処理部21-1は、図8の行Cに示すように現フレームの音声波形の先頭に付加した1ピッチ長 $L_p$ の区間に、行Dに示すように前記波形Zを挿入する。

図8、行Aでは、波形Xは波形伸張用バッファ23内の1フレーム前の信号から切り出し、波形Yは波形処理用バッファ21-0内の現フレームの信号から切り出しているが、ピッチ長 $L_p$ がフレーム長 $L_f$ の1/2以下の場合には、現フレームの音声波形のみを使って挿入波形を作成してもよい。例えば、図9、行A中に示すように、図8、行Aの波形X、波形Yに対応する波形X'、Y'を現フレームの先頭から連続して2ピッチ長の区間から切り出し、切り出した波形X'、Y'に三角窓をかけて互いに加算して波形Z'（図9、行B）を作成し、現フレームの波形X'とY'の間（図9、行C）に波形Z'を挿入して長さ $L_f+L_p$ の波形とすれば（図9、行D）、図8の手法と同様の効果が得られる。

- [0031] 図9の手法では、図8の手法で使用した波形伸張用バッファ23は不要になるメリットがある反面、ピッチ長 $L_p$ がフレーム長 $L_f$ の1/2以下である制限がある。しかしながら、波形処理用バッファ21-0のサイズを例えば2フレーム分として現フレームとその直前のフレームの波形を保持するようにし、伸張処理を2フレーム分の音声信号に対して行うことにより、ピッチ長 $L_p$ が $L_f/2$ を超えても波形伸張処理を行うことができる。

フレーム波形伸張部21の処理によって、長さ $L_f$ の現フレームの音声信号波形は、図8、行D及び図9、行Dに示すように長さが $L_f+L_p$ に伸張された信号波形となって出

力される。

- [0032] 図3におけるフレーム波形短縮部22の機能構成例を図11に、フレーム波形短縮部22による波形短縮処理の例を図12に示す。フレーム波形短縮部22は、波形処理用バッファ22-0と、波形置換処理部22-1と、第3波形切出し部22-2と、ピッチ波形生成部22-3と、第4波形切出し部22-4とから構成されている。

第3波形切出し部22-2は、図12の行Aに示すように、波形処理用バッファ22-0に保持されている現フレームの音声信号波形の先頭サンプルから1ピッチ長 $L_p$ の区間の波形Dを切り出す。第4波形切出し部22-4は現フレームの音声信号波形の波形Dに続く1ピッチ長 $L_p$ の区間の波形Eを切出す。

- [0033] ピッチ波形生成部22-3は、前記切り出した波形Dと波形Eにそれぞれ三角窓の重み付けを行った後加算して、図12の行Bに示す波形Fを作成する。ここで利用する重み付け三角窓としては、やはり非特許文献1に記載されているが、波形区間の始点から終点に向かって、波形Dでは1から0に、波形Eでは0から1に直線的に変化する形状を使用できる。

波形置換処理部22-1は、行Aに示す現フレームの連続する計2ピッチ長の波形D、Eの区間を行Cに示すように1ピッチ長の区間に短縮して、1ピッチ長の前記波形Fで置換する(行D)。

- [0034] フレーム波形短縮部22の処理によって、フレーム長 $L_f$ の入力音声信号は、 $L_f - L_p$ に短縮された信号となって出力される。

フレーム波形短縮部22において、ピッチ長 $L_p$ がフレーム長 $L_f$ の1/2を超えることがある場合は、前記短縮処理を行なうことができない。つまり波形Eがフレーム内から切り出せないためである。例えば、フレーム長 $L_f$ を20ミリ秒とすると、ピッチ長 $L_p$ は10ミリ秒以下でなくてはならず、これはピッチ周波数が100Hz以上でなくてはならないことを意味する。男性音声では、ピッチ周波数が100Hz未満である場合もある。このようにピッチ長 $L_p$ がフレーム長 $L_f$ の1/2を超える場合は、フレーム波形短縮部22の波形処理用バッファ22-0のサイズを2フレーム分とし、現フレームとその直前のフレームの連続する2フレーム長の音声信号に対し、上述の短縮処理を行えばよい。

- [0035] ピッチ長 $L_p$ がフレーム長 $L_f$ よりも長い場合には、前記2フレーム分の入力音声信号



に対する短縮処理も、前記フレーム波形伸張部21の処理も、ともに行なうことができないが、ピッチ長が20ミリ秒より大きい、即ちピッチ周波数が50Hzよりも低いことは減多にありえないことから、仮に入力されたピッチ長 $L_p$ がフレーム長 $L_f$ よりも長い場合には、フレーム波形伸張処理、フレーム波形短縮処理とも無処理のまま入力信号をそのまま出力すればよい。

再び図3に戻って、制御部16は状態検出部14からの最大遅延ゆらぎ $J$ と受信バッファ12の蓄積パケット数(バッファ量) $S$ とに基づいて、受信バッファ12に蓄積するパケットの量を増やすべきか、減らすべきか、そのまま維持するのかの判断を行う。

[0036] 蓄積するパケットの量を増加すべきと判断する状況の例としては次の場合が考えられる。

- (a) 受信バッファに蓄積されている音声パケットの量が減少傾向にある場合。
- (b) 受信バッファに蓄積されている音声パケットの数が所定値より少なくなった場合。
- 。
- (c) 受信パケットの到着時間の間隔が増加傾向にある場合。

蓄積するパケットの量を減少すべきと判断する状況の例としては次の場合が考えられる。

- [0037] (a) 受信バッファに蓄積されている音声パケットの量が増加傾向にある場合。
- (b) 受信バッファに蓄積されている音声パケットの数が所定値以上に達した場合。
- (c) 受信パケットの到着時間の間隔が短縮傾向にある場合。

それ以外で、パケット到着時点の最大遅延ゆらぎに対して、その時点で蓄積しているパケット量が適切であると判断されるときは、そのまま維持という判断をする。尚、実際の遅延ゆらぎが何ミリ秒の時にパケット蓄積量はいくらが適切であるかについては図4に例を示したが、一定期間(例えば2秒間)内のパケット受信部11へのパケット到着間隔の最大値よりも、蓄積されたパケットの数 $S$ で算出されるフレームの長さの合計時間 $S \times L_f$ が少し長い程度がよいと考えられる。

[0038] 制御部16は、蓄積するパケットの量を増やすべきであると判断した場合、スイッチ $SW_1$ 、 $SW_2$ をそれぞれ端子 $A_1$ 、 $A_2$ に切り替える。蓄積するパケットの量を減らすべきであると判断した場合には、スイッチ $SW_1$ 、 $SW_2$ をそれぞれ端子 $C_1$ 、 $C_2$ に切り替

える。蓄積するパケットの量をそのまま維持すべきであると判断した場合には、スイッチSW1、SW2をそれぞれ端子B1、B2に切り替え、各切り替え位置により消費量調整部20の消費量が設定される。

波形伸張用バッファ23は、スイッチSW2の出力側の音声データ列を蓄積し、蓄積された音声データ列は、前述のように、フレーム波形伸張部21で利用される。

[0039] 音声データ列がサウンドデバイス18に送られると、所定速度のクロックに同期して出力音声は再生される。サウンドデバイス18は受け取った時間長の音声データ列の再生が終わると、次のパケットから復号した音声データ列を受け取る。

前述のように、サウンドデバイス18の中にもサウンドデバイス用のバッファを持つのが一般的で、ダブルバッファと呼ばれる方法がよく用いられる。ダブルバッファの両方のバッファが満杯のときは、一方のバッファの再生が終了し、そのバッファが空になるまでは次の音声データ列は受け取らない。

[0040] スイッチSW1、SW2がそれぞれ端子C1、C2側、即ちフレーム波形短縮部22側にセットされると、もとの1フレーム長に満たない長さの信号がスイッチSW2より出力される。1フレーム長に満たない長さの信号をそのままサウンドデバイス18に送ると、サウンドデバイスのオーバーヘッドが大きくなって、時に音声の途切れが発生することがある。サウンドデバイス18の仕様としてフレーム長の許容最低値がある場合には、スイッチSW2とサウンドデバイス18の間に中間のバッファを設けるのがよい。

スイッチSW1、SW2を端子A1、A2側に切り替ええると、音声パケット復号部13から出力された復号音声データ列は、フレーム波形伸張部21を通過してサウンドデバイス18に送られる。長さ $L_f$ の復号音声データ列がフレーム波形伸張部21を通過することによって、長さ $L_f + L_p$ のデータ列に伸張されるので、サウンドデバイス18での再生時間も $L_f + L_p$ になる。即ち、通常、サウンドデバイス18が音声データ列を受け取る時間間隔が $L_f$ であったのが、 $L_f + L_p$ の信号の再生時には音声データ列を受け取る時間間隔は $L_f + L_p$ になる。

[0041] 音声パケット復号部13は、サウンドデバイス18が音声データ列を受け取らないと受信バッファ12に対して次の送信要求を出さないなので、パケット受信部11が一定の間隔でパケットを受信していれば、受信バッファ12に蓄積されるパケット量は平均として

増えることになる。平均として増えるとは、 $L_p < L_f$ であるので、1回のフレーム波形伸張処理では、受信バッファに蓄積されるパケット量の増分は1フレーム分には満たない。復号音声信号に対し複数フレーム(Nとする)にわたってフレーム波形伸張部21の伸張処理を行うと、Nよりも小さいMフレーム分だけ受信バッファ12に蓄積されるパケット量が増加する。

[0042] 図13を用いてその様子を説明する。図13の行Aは受信バッファ12に順次蓄積されるパケット $P_1, P_2, P_3, \dots$ を、また図13の行Bは定常の音声フレーム $F_1, F_2, F_3, \dots$ の消費状態を示す。つまり、スイッチSW1とSW2が端子B1、B2にセットされている場合は受信バッファ12に格納されているパケットを1個ずつ取り出して音声パケット復号部13で1フレーム長 $L_f$ 毎の音声データ列に復号し、圧縮、伸張の処理を受けずにスルーパス24を通してサウンドデバイス18に与えられる。サウンドデバイス18では1フレーム長 $L_f$ 毎の音声データ列を音声に再生する。従って、この状態では1パケットを消費する毎に1フレーム長の音声再生され、1パケット分の音声再生されると、次のパケットが受信バッファ12から取り出され、音声パケット復号部13で音声データ列の復号が行われる。

[0043] これに対し、図13の行Cに示す消費量小の状態では、スイッチSW1とSW2は端子A1とA2に接続され、復号音声データ列は伸張処理を受けるので、サウンドデバイス18に与えられる音声データ列の時間長は $L_f + L_p$ となる。従って、行Cに示すように伸張されたフレーム $F'_1, F'_2, F'_3, \dots$ がサウンドデバイス18に与えられることになる。サウンドデバイス18は1パケットから復号した各フレーム長 $L_f$ の音声データ列を $L_f + L_p$ の時間を掛けて音声を再生することになるからサウンドデバイス18側から音声パケット復号部13に復号音声データ列を出力することを要求するタイミングは $L_f + L_p$ の周期となる。図13、行Bに示した定常消費状態の例では時間 $T_M$ 内に6個のパケットを消費したが、図13、行Cに示す受信バッファ12からパケットを取り出す周期が $L_f + L_p$ の消費状態例ではパケットを4個消費することになる。これによりパケットの消費量を定常の消費状態より少なくすることができる。

[0044] スイッチSW1、SW2を端子C1、C2側に切り替えると、音声パケット復号部13から出力された復号音声データ列は、フレーム波形短縮部22を通してサウンドデバイス

18に送られる。長さ $L_f$ の復号音声データ列がフレーム波形短縮部22を通ることによって、長さ $L_f - L_p$ の音声データ列に短縮されるので、図13、行Dに示すように短縮されたフレーム $F''_1, F''_2, F''_3, \dots$ がサウンドデバイス18に与えられることになる。そのため、サウンドデバイス18での各短縮フレームの再生時間も $L_f - L_p$ になる。即ち、通常、サウンドデバイス18が音声データ列を受け取る時間間隔が $L$ であったのが、 $L_f - L_p$ の信号の再生時には間隔が $L_f - L_p$ になる。この結果、音声 packets 復号部13も、通常1フレーム長 $L_f$ の時間間隔よりも短い間隔で受信バッファ12に対して次の送信要求を出すので、packet 受信部11が一定の間隔でpacketを受信していれば、受信バッファに蓄積されるpacket量は平均として減ることになる。平均として減るとは、 $L_p < L_f$ であるので、1回のフレーム波形短縮処理では、受信バッファに蓄積されるpacket量の減少分は1フレーム分には満たない。複数フレーム( $N$ とする)にわたってフレーム波形短縮部22を通ると、 $N$ よりも小さい $M$ フレーム分受信バッファに蓄積されるpacket量が減少する。

[0045] 図13の行Dに示す例では、フレーム短縮部22を通すことにより、同じ時間 $T_M$ の間にpacketを10個消費することを表わしている。このようにpacketの消費量を大きくすることにより、受信バッファ12に蓄えられるpacketの数を少なくすることができる。

制御部16により、より高度なバッファ量制御が可能である。例えば、状態検出部14からのバッファ量 $S$ と最大遅延ゆらぎ $J$ に基づいて、受信バッファに蓄積するpacketの量を増やすべき、減らすべきと判断する際に、緊急に増やすべき／ゆるやかに増やすべき、緊急に減らすべき／ゆるやかに減らすべき、という増減速度を判断に加えることができる。具体的には、通信網の状態が突然悪化したときに、受信バッファに蓄積するpacketの量をゆるやかに増やしていたのでは、音声の途切れが発生してしまうかもしれない。通信網の状態変化が急激であれば、バッファに蓄積するpacket量の制御も緊急に行うべきである。逆に、一般にドリフトと言われる、送信側と受信側のクロックずれやタイミングのわずかなずれの蓄積により、バッファに蓄積するpacket量が徐々に所望の量より増えてきてしまった、あるいは減ってきてしまったという場合には、ゆるやかに増減すればよい。

[0046] 緊急に増減が必要な場合は、音声区間、非音声区間にかかわらず、消費量調整

部20のスイッチSW1、SW2をそれぞれ端子A1、A2側、あるいは端子C1、C2側にセットして、迅速な波形伸縮を行うことによって、パケット蓄積量の増減も迅速に制御することができる。逆に、ゆるやかな増減でよい場合は、音声分析部15における音声／非音声判定の結果、非音声フレームにおいてのみスイッチSW1、SW2をそれぞれ端子A1、A2側、あるいは端子C1、C2側にセットするという方法でもよい。さらに、非音声フレームの場合には、ピッチ長を実際にピッチ分析した値ではなく、フレーム長L<sub>f</sub>の1/2以下の任意の値に設定してよいので、よりゆるやかな増減の場合には、任意に設定するピッチ長を短い値にするとよい。

- [0047] 図14は図3の構成において図6に示した音声分析部15による音声／非音声判定結果と、状態検出部14により検出した現受信パケットに対し過去T秒間(例えば2秒間)における最大遅延ゆらぎJと、受信バッファ12内の蓄積パケット数Sとに基づいて、バッファ量の調整が緊急を要するかその緊急度を判定する手順の例を示す。
- ステップS1:状態検出部14における最大遅延ゆらぎJを受信バッファ12内の各パケットの受信時刻から求める。
- ステップS2:制御部16により最大遅延ゆらぎJに対応する最適バッファ量Bを図4の表16Tを参照して決める。
- ステップS3:状態検出部14により現時点における受信バッファ12のバッファ量(蓄積されているパケット数)Bを求める。
- ステップS4:制御部16により決定した最適バッファ量Bと状態検出部14により検出した実際のバッファ量Sとの差の絶対値  $|S - B|$  が予め決めた正の値Eより小であるか判定し、小であればバッファ量の調整は必要ないものと判断し、現状を維持する(これを緊急度0と定義する)。
- ステップS5:差の絶対値  $|S - B|$  がEより小でない場合は、バッファ量の調整が必要であることを意味し、差S-Bが-E以下であるが判定する。-E以下であることはバッファ量を増大する必要があることを意味し、以下のステップS6, S7を実行してバッファ量を増大する緊急度を判定する。S-Bが-E以下でない場合はS-BがE以上であり、バッファ量を減少する必要があることを意味し、以下のステップS8, S9を実行してバッファ量を減少する緊急度を判定する。

ステップS6:現時点のバッファ量Sが0以上でかつ最適バッファ量Bの20%以下であるか判定し、もしそうであればバッファ量の調整(ここでは増大)の緊急度が大きいと判定する。なお、Bが1以上でSが0のとき、即ちバッファが枯渇して音切れの危険がある状態もこのステップで緊急度が大きいと判定される。

ステップS7:現時点のバッファ量Sが最適バッファ量Bの20%より大で、かつ50%以下であるか判定する。そうであればバッファ量調整の緊急度は中と判定し、そうでなければ緊急度は小と判定する。

ステップS8: $S - B < -E$ の場合はバッファ量を増加する必要がある、現時点のバッファ量Sが最適バッファ量Bの200%以上であるか判定する。もしそうであれば、緊急度大と判定する。

ステップS9:SがBの200%以上でない場合は、SがBの200%より小で、かつ150%以上であるか判定し、もしそうであれば、緊急度は中であると判定し、そうでなければ緊急度は小であると判定する。

#### 制御例1

図15に示す表1は、図14の手順により判定した緊急度(大、中、小、0)と、音声／非音声判定結果に基づいて制御部16が実行する消費量調整部20の第1の制御例を示す。

[0048] 判定結果が緊急度大の場合、音声信号の音声／非音声判定結果にかかわらず、スイッチSW1, SW2を、バッファ量を増すべきときは端子A1, A2側に、バッファ量を減らすべきときは端子C1, C2側に制御して現フレームの復号音声データ列に対し、伸張／短縮処理を必ず行う。判定結果が緊急度中の場合も緊急度大の場合と同じ制御を行う。

判定結果が緊急度小の場合は、現フレームの復号音声データ列が音声区間であればスイッチSW1, SW2を端子B1, B2に固定して伸張／短縮を行わず、非音声区間であれば、スイッチSW1, SW2を、バッファ量を増すべきときは端子A1, A2側に、バッファ量を減らすべきときはC1, C2側にセットする。

#### 制御例2

図16に示す表2は、図14の手順により判定した緊急度(大、中、小、0)と、音声／

非音声判定結果に基づいて制御部16が実行する消費量調整部20の第2の制御例を示す。この制御例は第1の制御例に比べて、緊急度が小のときの音声区間でN1フレームに1回の割合でA1, A2側又はC1, C2側にスイッチSW1, SW2をセットし、それ以外のフレームではB1, B2側にセットしている。これは音声の状態(通話環境)によっては、非音声区間がまったく検出されない場合が想定されるため、非音声区間が検出されない場合でも、バッファ量を変更できるようにしたものである。N1の値は1以上の整数であるが、例えばN1=5とする。その他は第1の制御例と同様である。

### 制御例3

図17に示す表3は、制御部16が実行する消費量調整部20の第3の制御例を示す。この例は、図16に示した第2の制御例における緊急度が中の制御を緊急度が中の制御と異ならしており、音声区間ではN2フレームに1回の割合でA1, A2側又はC1, C2側にスイッチSW1, SW2をセットし、それ以外のフレームではB1, B2側にセットする。N2の値は1以上の整数であるが、N1より小さい(即ちA1, A2側又はC1, C2側にセットする頻度が制御例2の対応する制御より小さい)例えばN2=2とするのが好ましい。その他は第2の制御例と同様である。

### 制御例4

図18に示す表4は、制御部16が実行する消費量調整部20の第4の制御例を示す。この制御例では、緊急度と音声／非音声区間のすべての組に対応してそれぞれ予め決めたフレーム数ごとに1回の割合でスイッチSW1, SW2をA1, A2側又はC1, C2側にセットし、それ以外ではB1, B2側にセットする汎用的な手法を定義している。即ち、緊急度が小と大の場合の音声区間での制御は図17の第3制御例と同様であるが、さらに緊急度が大の場合には音声区間／非音声区間にかかわらずN3フレームごとに1回の割合でスイッチSW1, SW2をA1, A2側又はC1, C2側にセットする。また、緊急度が中の場合には、非音声区間においてN4フレームごとに1回の割合でスイッチSW1, SW2をA1, A2側又はC1, C2側にセットし、それ以外のフレームではB1, B2側にセットし、緊急度が小の場合には、非音声区間においてN5フレームごとに1回の割合でスイッチSW1, SW2をA1, A2側又はC1, C2側にセットし、それ以外のフレームではB1, B2側にセットする。

- [0049] N3の値は1以上の整数であり、例えばN3=1とする。N4の値は1以上の整数であり、例えばN4=1とする。N5の値は1以上の整数であり、例えばN5=2とする。これらの値N1～N5を適切に選ぶことにより、音質の低下(違和感の増加)とバッファ量の変更速度のバランスを調整することができる。

## 第2実施例

図6に示した音声分析部15では、固定の閾値PthとフレームパワーPを比較することにより現フレームの復号音声信号が音声区間であるか非音声区間であるかの判定を行っている。この構成は簡便でよいが、復号音声信号にレベルの高い背景雑音(空調騒音やオフィスのざわつき、街頭騒音など)が含まれている場合に、音声区間／非音声区間を正確に判定できない場合がある。音声／非音声判定に背景雑音を考慮することによりこの点を改善した構成を図19に示す。

- [0050] 図19に示す音声分析部15の構成は、図6の構成に対しさらに動的パワー閾値計算部150が設けられ、ピッチ相関計算部158においてはピッチ分析を行ってピッチ長 $L_p$ を求めると共にピッチ相関値(相関ピーク値) $r$ を求める。動的パワー閾値計算部150は所定時間(例えば数秒乃至数十秒)過去に遡った範囲内のそれぞれのフレームのパワー $P_f$ の系列とそれぞれのフレームについて求めたピッチ相関値 $r$ を用いてパワー閾値 $P_d$ を動的に決定する。例えば、ピッチ相関 $r$ が高いフレーム、例えば $r \geq 0.7$ のフレームはほぼ確実に音声区間であるとみなしてそれらのフレームのパワーの長時間平均 $P_{av1}$ を計算することにより音声信号レベルが推定される。
- [0051] 一方、ピッチ相関 $r$ が低いフレーム、例えば $r \leq 0.4$ のフレームは、音声区間でない(即ち、背景雑音)か、音声区間の無声音区間(有声音区間でない)と推定される。背景雑音か無声音かの正確な区別は難しいが、無声音区間は相対的に少ないとみなせば、当該背景雑音か無声音区間とみなされるフレームのパワーの長時間平均 $P_{av2}$ を計算することによって、非音声区間の信号レベルが推定される。もちろん、背景雑音と無声音を区別して、背景雑音のフレームのみでパワーの長時間平均を計算できるほうが好ましいので、ピッチ相関値 $r$ が低いフレームでも、フレームパワーの時系列の定常性を観測して、定常性の低い区間は無声音区間とみなし、上記パワーの長時間平均の計算から除外してもよい。



[0052] 動的パワー閾値計算部150では、推定された音声信号レベル $P_{av1}$ と、非音声区間(即ち、背景雑音区間)の信号レベル $P_{av2}$ に基づいて、動的パワー閾値 $P_d$ を動的に決定、更新する。例えば、動的パワー閾値 $P_d$ はレベル $P_{av1}$ と $P_{av2}$ の間の値に選ぶが、好ましくは非音声区間のレベル値 $P_{av2}$ よりわずかに大きい値に決めるのがよい。区間判定部154は、フレームのパワー $P_f$ が動的パワー閾値 $P_d$ よりも大きければそのフレームを音声区間と判定し、そうでなければ非音声区間と判定する。

上述の第2実施例による判定結果は、前述の図15～18に示した第1乃至第4制御例のいずれに適用してもよい。

### 第3実施例

前述の第1及び第2実施例では各フレームが音声区間か非音声区間かの2通りに判定していたが、この実施例では音声区間が有声音区間であるか無声音区間であるか区別し、非音声区間が背景雑音区間であるか無音区間であるか区別する。従って、各フレームを、有声音、無声音、背景雑音、無音の4通りに判定し、その判定に基づいて消費量調整部20を制御する。そのための音声分析部15の構成を図20に示す。

[0053] 図20に示す音声分析部15の構成は、図19に示した構成において区間判定部154にもピッチ相関値 $r$ を与え、区間判定部154においてパワー閾値として動的パワー閾値 $P_d$ と固定パワー閾値 $P_{th}$ の両方を使用し、フレームパワー $P_f$ とピッチ相関値 $r$ からフレームが上記4通りのいずれの区間かを判定する。

図21は区間判定の処理手順を示す。

ステップS1:フレームのパワー $P_f$ が固定閾値 $P_{th}$ 以下か判定し、そうであれば無音区間と判定する。

ステップS2: $P_f$ が $P_{th}$ 以下でない場合は、フレームパワー $P_f$ が動的パワー閾値 $P_d$ 以下か判定し、そうであれば背景雑音区間と判定する。

ステップS3: $P_f$ が $P_d$ 以下でない場合は、ピッチ相関値 $r$ が予め決めた正の値 $R_c$ 以下か判定し、そうであれば音声区間の無声音区間と判定し、そうでなければ音声区間の有声音区間と判定する。

[0054] 図22に示す表5は、図14の手順により判定した緊急度(大、中、小、0)と、図21の

処理による区間判定結果に基づいて制御部16が実行する消費量調整部20の制御例を示す。この制御例では、緊急度が中と判定された場合は、区間判定結果に関係なく無条件で消費量調整部20のスイッチSW1, SW2を端子A1, A2側又はC1, C2側にセットする。

緊急度が中と判定された場合は、判定結果が有声音、無声音、背景雑音、無音、に対しそれぞれ予め決めた整数値N6, N7, N8, N9のフレーム数ごとに1フレームの割合でスイッチSW1, SW2を端子A1, A2側又はC1, C2側にセットする。予め決めた整数値としては例えばN6=2, N7=2, N8=1, N9=1を使用することができるが、これらに限定されるものでない。

- [0055] 緊急度が小と判定された場合も同様に、判定結果が有声音、無声音、背景雑音、無音、に対しそれぞれ決めた整数値N10, N11, N12, N13のフレーム数ごとに1フレームの割合でスイッチSW1, SW2を端子A1, A2側又はC1, C2側にセットする。予め決めた整数値としては例えばN10=5, N11=4, N12=4, N13=2を使用することができるが、これらに限定されるものでない。

整数値N6～N13を適切に選択することにより、音質の低下(違和感の増加)とバッファ量の変化速度のバランスを調整することができる。

- [0056] 尚、上述の各実施例ではサウンドデバイス18を終段に接続した実施例を説明したが、受信した音声データ列を蓄積するだけの場合もあり、必ずしも音声を再生しない場合もある。

以上説明した本発明の音声パケット再生方法はコンピュータに本発明による音声パケット再生プログラムを実行させることにより実現することができ、またコンピュータに本発明による音声パケット再生プログラムをインストールし、CPUに解読させて実行させることにより、コンピュータによって本発明の音声パケット再生装置を構築することができる。本発明による音声パケット再生プログラムはコンピュータが解読可能なプログラム言語によって記述され、コンピュータが読み取り可能な例えば磁気ディスク或はCD-ROMのような記録媒体に記録され、これらの記録媒体からコンピュータにインストールするか、或は通信回線を通じてコンピュータにインストールされ、コンピュータに備えられたCPUに解読されて音声パケット再生動作を実行する。

### 産業上の利用可能性

[0057] IP通信網上で音声通信を行う利用形態が普及してきており、本発明を適用することによって、安価で信頼性の高い音声通信が実現できる。

### 請求の範囲

- [1] 入力音声データ列のフレーム毎の符号化により生成された音声符号を収容して送信された一連の音声パケットを受信して再生する音声パケットの再生方法であり、以下のステップを含む：
- (a) 受信パケットを受信バッファに蓄積し、
  - (b) 上記受信パケットの到着時間のゆらぎを一定期間観測して得られるゆらぎの最大値又は統計値、以下最大遅延ゆらぎと呼ぶ、と上記受信バッファに蓄積されているパケット数であるバッファ蓄積パケット数とを検出し、
  - (c) 上記最大遅延ゆらぎから、予め決めた最大遅延ゆらぎと最適蓄積パケット数の関係を用いて、上記受信バッファに蓄積すべき最適なパケット数、以下最適蓄積パケット数と呼ぶ、を求め、
  - (d) 上記検出したバッファ蓄積パケット数と上記最適蓄積パケット数の相違の程度を複数の段階で判定し、
  - (e) 上記受信バッファから現フレームに対応するパケットを取り出し、そのパケット中の音声符号を復号して現フレームの復号音声データ列を得て、
  - (f) 上記相違の程度の段階ごとに定めた規則に基づいて、上記バッファ蓄積パケット数を上記最適蓄積パケット数に近づけるように、上記復号音声データ列に対しその波形の伸張、短縮又は現状維持のいずれかの処理を行い、現フレームの再生音声データとして出力する。
- [2] 請求項1記載の音声パケット再生方法において、上記ステップ(f) は、
- (f-1) 上記復号音声データ列についてのピッチ長を求めるステップと、
  - (f-2) 上記音声データ列を分析して音声区間であるか非音声区間であるかを判定する音声非音声判定ステップと、
  - (f-3) 上記音声非音声判定結果と、上記相違の程度の段階の判定結果とに基づいて、上記復号音声データ列に対し上記ピッチ長に対応する波形を挿入又は削除するか又は上記復号音声データ列を変更しないことにより伸張、短縮又は現状維持の処理を行うステップ、
- とを含む。

- [3] 請求項2記載の音声パケット再生方法において、  
上記ステップ(d)は、上記相違の程度の段階が、上記バッファ蓄積パケット数を緊急に増加又は減少させるべき状況である緊急度大であるか、上記バッファ蓄積パケット数を緩やかに増加又は減少させるべき状況である緊急度小であるかを判定し、  
上記ステップ(f-3)は、上記段階が緊急度大の場合は、上記音声区間又は非音声区間にかかわらず上記復号音声データ列の波形の伸張又は短縮を実行し、上記段階が緊急度小の場合は、上記復号音声データ列が非音声区間の場合にのみ上記復号音声データ列の波形の伸張又は短縮を実行するステップを含む。
- [4] 請求項2記載の音声パケット再生方法において、  
上記ステップ(d)は、上記相違の程度の段階が、上記バッファ蓄積パケット数を緊急に増加又は減少させるべき状況である緊急度大であるか、上記バッファ蓄積パケット数を緩やかに増加又は減少させるべき状況である緊急度小であるかを判定し、  
上記ステップ(f-3)は、上記段階が緊急度大の場合は、上記音声区間又は非音声区間にかかわらず上記復号音声データ列の波形の伸張又は短縮を実行し、上記段階が緊急度小の場合は、上記復号音声データ列が音声区間の場合には予め決めたフレーム数N1毎に1回の割合で上記復号音声データ列の波形の伸張又は短縮を実行し、上記復号音声データ列が非音声区間の場合には予め決めたフレーム数N2ごとに1回の割合で上記音声復号データ列の波形の伸張又は短縮を実行するステップを含み、上記N1及びN2は1以上の整数であり、かつN2はN1より小である。
- [5] 請求項1記載の音声パケット再生方法において、上記ステップ(f)は、  
(f-1) 上記復号音声データ列についてのピッチ長を求めるステップと、  
(f-2) 上記復号音声データ列を分析して有声音区間、無声音区間、背景雑音区間、無音区間のいずれであるか区間判定するステップと、  
(f-3) 上記区間判定結果と、上記相違の程度の段階の判定結果とに基づいて、上記復号音声データ列に対し上記ピッチ長に対応する波形を挿入又は削除するか又は上記復号音声データ列を変更しないことにより伸張、短縮又は現状維持の処理を行うステップ、  
とを含む。

- [6] 請求項5記載の音声パケット再生方法において、  
上記ステップ(d)は、上記相違の程度の段階が、上記バッファ蓄積パケット数を緊急に増加又は減少させるべき状況である緊急度大であるか、上記バッファ蓄積パケット数を緩やかに増加又は減少させるべき状況である緊急度小であるかを判定し、  
上記ステップ(f-3)は、上記段階が緊急度大の場合は、上記区間判定結果にかかわらず上記復号音声データ列の波形の伸張又は短縮を実行し、上記段階が緊急度小の場合は、上記有声音区間、上記無声音区間、上記背景雑音区間、上記無音区間に対しそれぞれ予め決めたフレーム数N1, N2, N3, N4毎に1回の割合で上記復号音声データ列の波形の伸張又は短縮を実行するステップを含み、上記N1, N2, N3, N4は正の整数であり、それらのうち少なくとも1つは2以上でかつ他の3つとは異なる値である。
- [7] 入力音声データ列のフレーム毎の符号化により生成された音声符号を収容して送信された一連の音声パケットを受信して再生する音声パケットの再生装置であり、  
音声パケットをパケット通信網から受信するパケット受信部と、  
受信したパケットを一時的に蓄積し、要求に応じてパケットを読み出す受信バッファと、  
上記受信パケットの到着時間のゆらぎを一定期間観測して得られるゆらぎの最大値又は統計値、以下最大遅延ゆらぎと呼ぶ、と上記受信バッファに蓄積されているパケット数であるバッファ蓄積パケット数を検出する状態検出部と、  
上記最大遅延ゆらぎから、予め決めた最大遅延ゆらぎと最適蓄積パケット数の関係を用いて、上記受信バッファに蓄積すべき最適なパケット数、以下最適蓄積パケット数と呼ぶ、を求め、上記検出したバッファ蓄積パケット数と上記最適パケット数の相違の程度を複数の段階で判定し、上記相違の程度の段階ごとに予め定めた規則に基づいて、上記バッファ蓄積パケット数を上記最適蓄積パケット数に近づけるように波形の伸張、短縮、又は現状維持を指示する制御信号を生成する制御部と、  
上記受信バッファから取り出された現フレームに対応するパケット中の音声符号を復号して現フレームの復号音声データ列を得る音声パケット復号部と、  
上記制御信号に従って、上記現フレームの復号音声データ列に対し、その波形の

伸張、短縮又は現状維持のいずれかの処理を行い、現フレームの再生音声データとして出力する消費量調整部、  
とを含む。

[8] 請求項7記載の音声パケット再生装置において、

上記音声分析部は上記復号音声データ列を分析して音声区間であるか非音声区間であるか判定してその判定結果を上記制御部に与えると共に、上記復号音声データ列についてのピッチ長を求めて上記消費量調整部に与えるようにされており、

上記制御部は上記判定結果と、上記相違の程度の段階の判定結果とに基づいて上記消費量調整部に対し、現フレームの上記復号音声データ列の伸張、短縮又は現状維持の制御を与えるようにされており、

上記消費量調整部は上記制御に従って上記ピッチ長に対応する波形を上記復号音声データ列に対し挿入又は削除するか又は上記復号音声データ列を変更しないようにされている。

[9] 請求項8記載の音声パケット再生装置において、

上記制御部は、上記相違の程度の段階が、上記バッファ蓄積パケット数を緊急に増加又は減少させるべき状況である緊急度大であるか、上記バッファ蓄積パケット数を緩やかに増加又は減少させるべき状況である緊急度小であるかを判定し、上記消費量調整部に対し、上記段階が緊急度大の場合は、上記音声区間又は非音声区間にかかわらず上記復号音声データ列の波形の伸張又は短縮を実行させる制御を与え、上記段階が緊急度小の場合は、上記復号音声データ列が非音声区間の場合にのみ上記復号音声データ列の波形の伸張又は短縮を実行させる制御を与えるようにされている。

[10] 請求項8記載の音声パケット再生装置において、

上記制御部は上記相違の程度の段階が、上記バッファ蓄積パケット数を緊急に増加又は減少させるべき状況である緊急度大であるか、上記バッファ蓄積パケット数を緩やかに増加又は減少させるべき状況である緊急度小であるかを判定し、上記段階が緊急度大の場合は、上記音声区間又は非音声区間にかかわらず上記消費量調整部に対し上記復号音声データ列の波形の伸張又は短縮を実行させる制御を与え

、上記段階が緊急度小の場合は、上記消費量調整部に対し、上記復号音声データ列が音声区間の場合には予め決めたフレーム数N1毎に1回の割合で上記復号音声データ列の波形の伸張又は短縮を実行させ、上記復号音声データ列が非音声区間の場合には予め決めたフレーム数N2ごとに1回の割合で上記音声復号データ列の波形の伸張又は短縮を実行させる制御を与えるようにされており、上記N1及びN2は1以上の整数であり、かつN2はN1より小である。

- [11] 請求項7記載の音声パケット再生装置において、上記音声分析部は上記復号音声データ列を分析して有声音区間、無声音区間、背景雑音区間、無音区間のいずれであるか区間判定してその判定結果を制御部に与えると共に、上記復号音声データ列についてのピッチ長を求めて上記消費量調整部に与えるようにされており、

上記制御部は、上記区間判定結果と、上記相違の程度の段階の判定結果とに基づいて、上記消費量調整部に対し、現フレームの上記復号音声データ列の伸張、短縮又は現状維持の処理を実行させる制御を与えるようにされており、

上記消費量調整部は、上記制御に従って上記ピッチ長に対応する波形を上記復号音声データ列に対し挿入又は削除するか又は上記復号音声データ列を変更しないようにされている。

- [12] 請求項11記載の音声パケット再生装置において、上記制御部は上記相違の程度の段階が、上記バッファ蓄積パケット数を緊急に増加又は減少させるべき状況である緊急度大であるか、上記バッファ蓄積パケット数を緩やかに増加又は減少させるべき状況である緊急度小であるかを判定し、上記消費量調整部に対し、上記段階が緊急度大の場合は上記区間判定結果にかかわらず上記復号音声データ列の波形の伸張又は短縮を実行させる制御を与え、上記段階が緊急度小の場合は、上記有声音区間、上記無声音区間、上記背景雑音区間、上記無音区間に対しそれぞれ予め決めたフレーム数N1, N2, N3, N4毎に1回の割合で上記復号音声データ列の波形の伸張又は短縮を実行させる制御を与えるようにされており、上記N1, N2, N3, N4のうち、少なくとも1つは2以上の整数であり、かつ他の3つとは異なる値である。

- [13] コンピュータが解読可能なプログラム言語によって記述され、コンピュータに請求項1記載の音声パケット再生方法を実施させる音声パケット再生プログラム。



- [14] コンピュータが読取り可能な記録媒体によって構成され、この記録媒体に請求項1  
3記載の音声パケット再生プログラムを記録した記録媒体。

[図1]

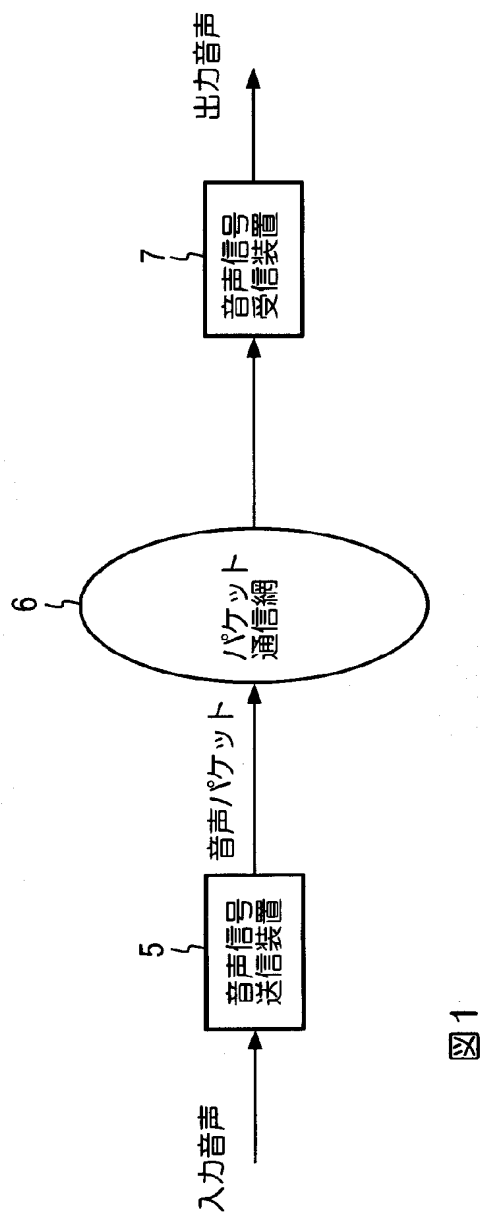


図1

[図2]

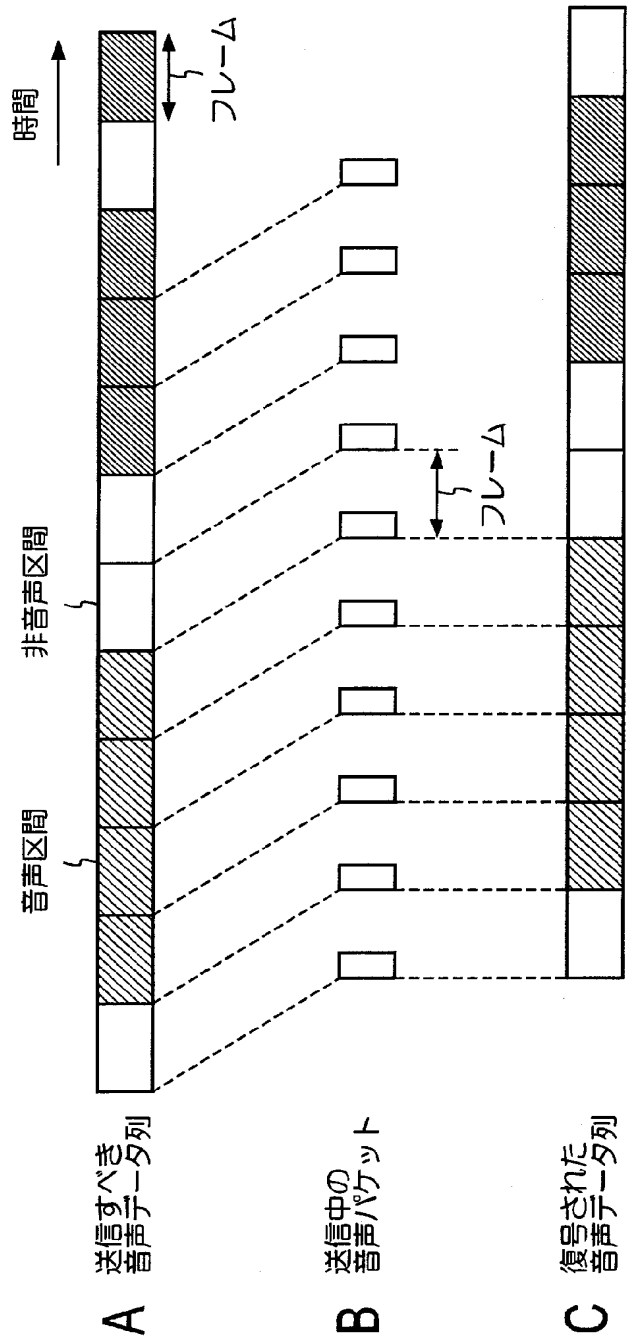


図2

[図3]

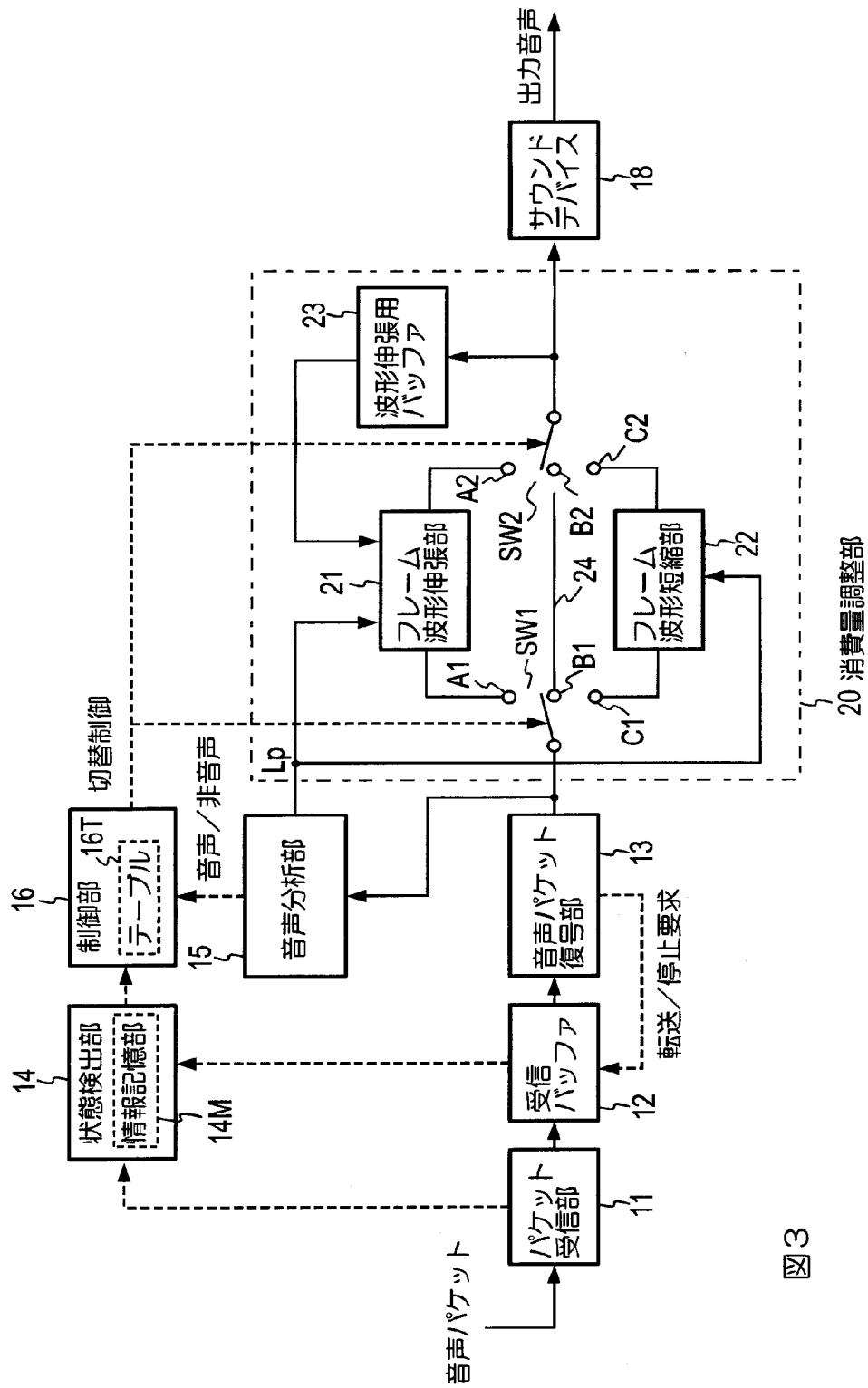


図3

[図4]

表 16T

最大遅延ゆらぎJ	バッファに蓄積すべき 最適パケット数B
15msまで	1
32msまで	2
48msまで	3
64msまで	4
80msまで	5
96msまで	6
128msまで	8
160msまで	10
192msまで	12
240msまで	15
320msまで	20
400msまで	25
480msまで	30
640msまで	40
800msまで	50

図4

[図5]

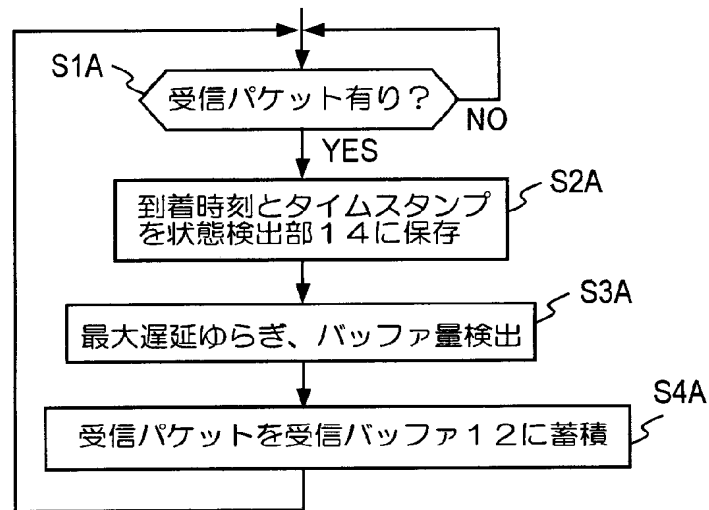


図5A

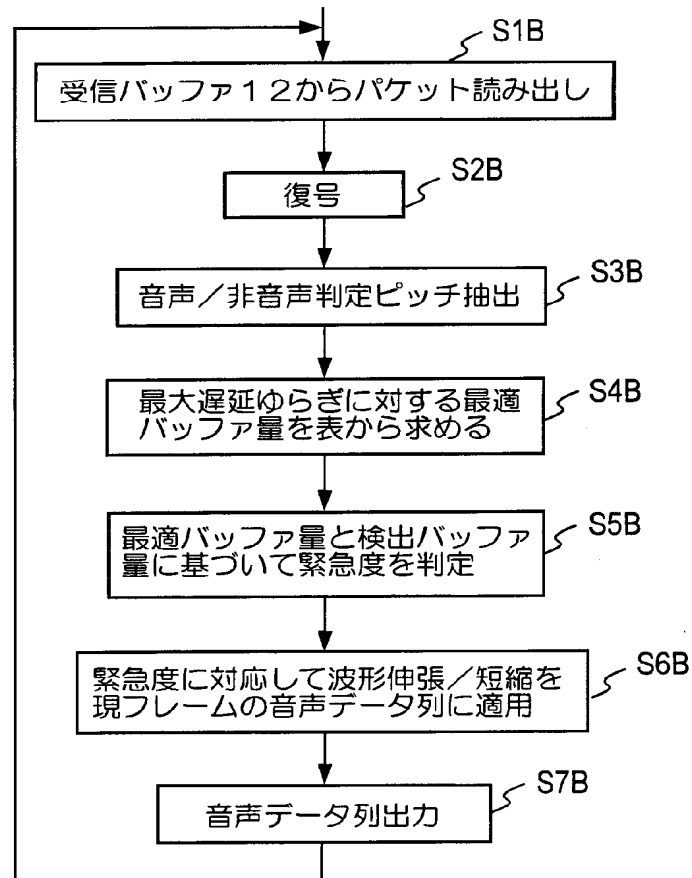


図5B

[図6]

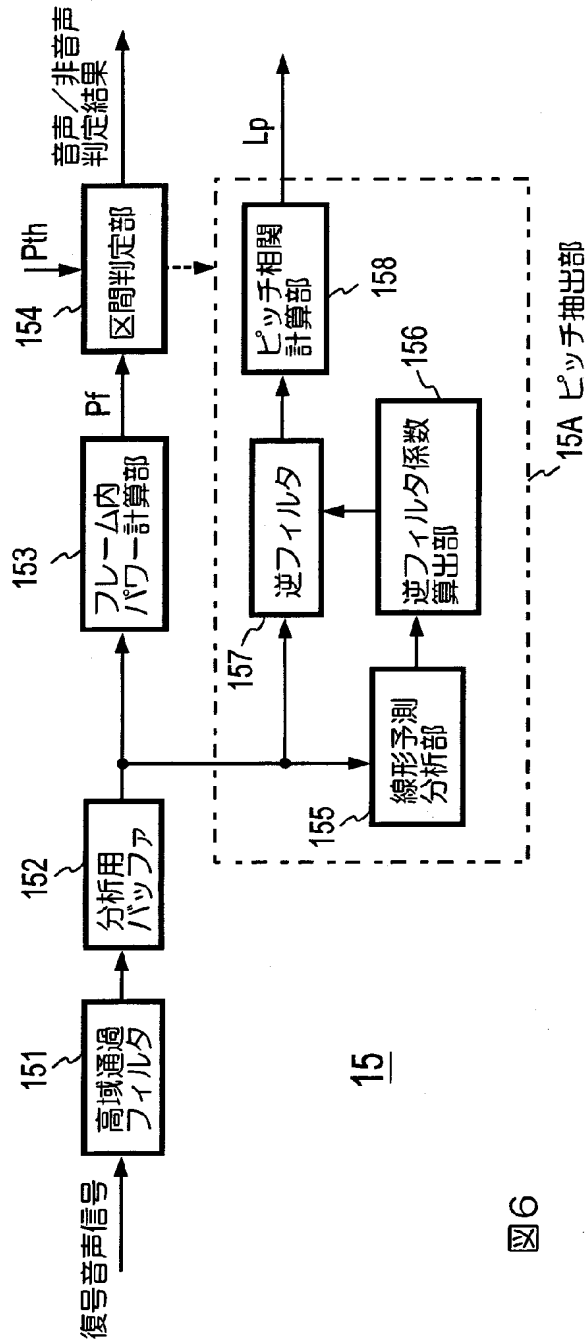


図6

[図7]

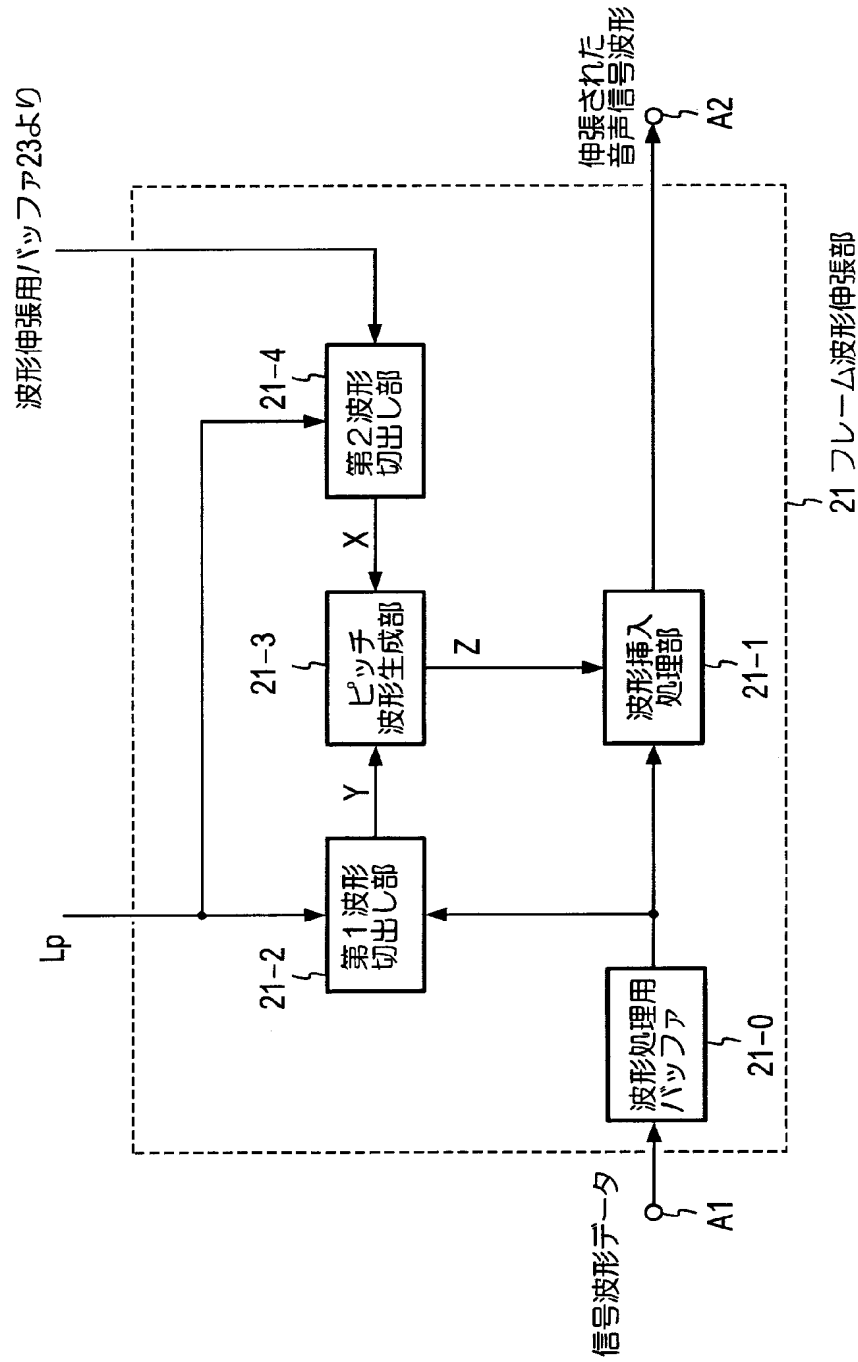
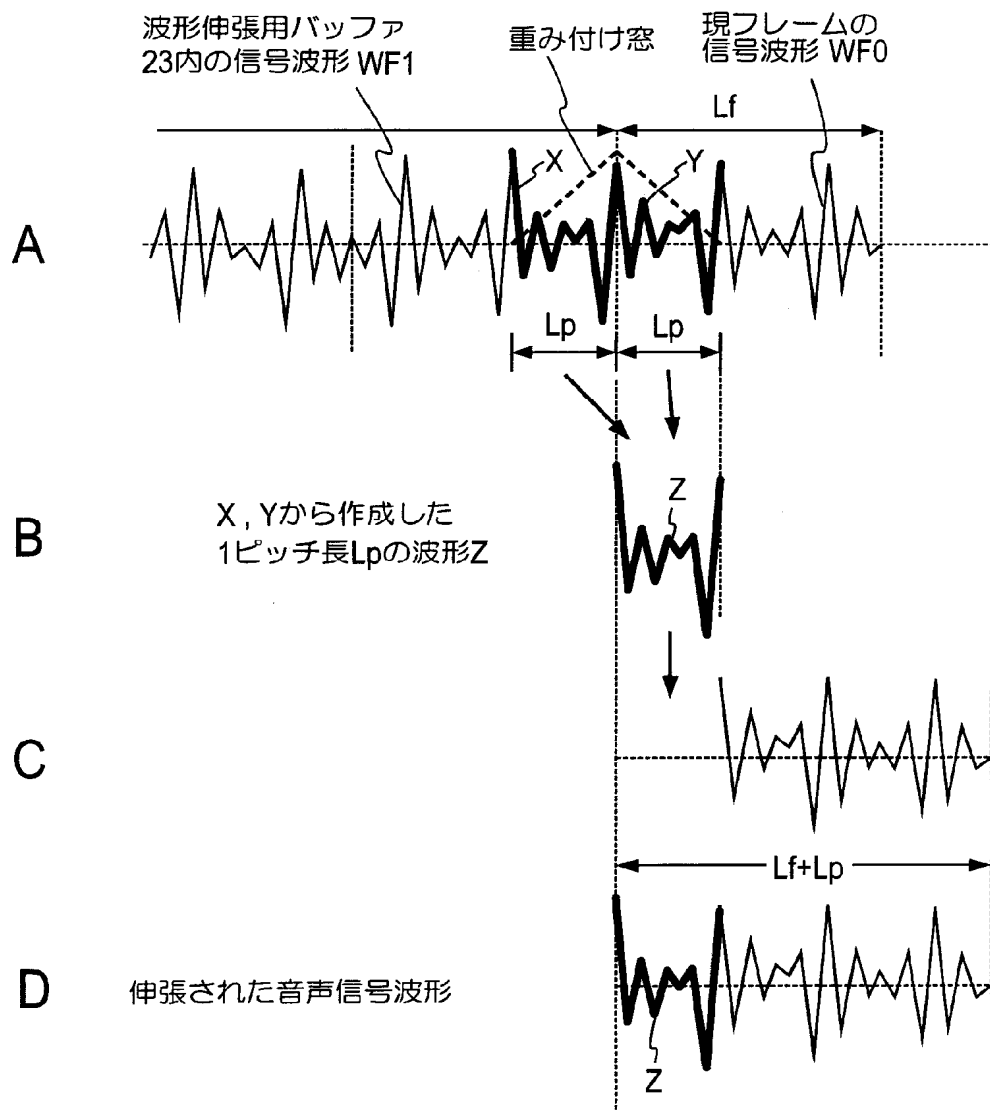


図 7



[図8]



[図9]

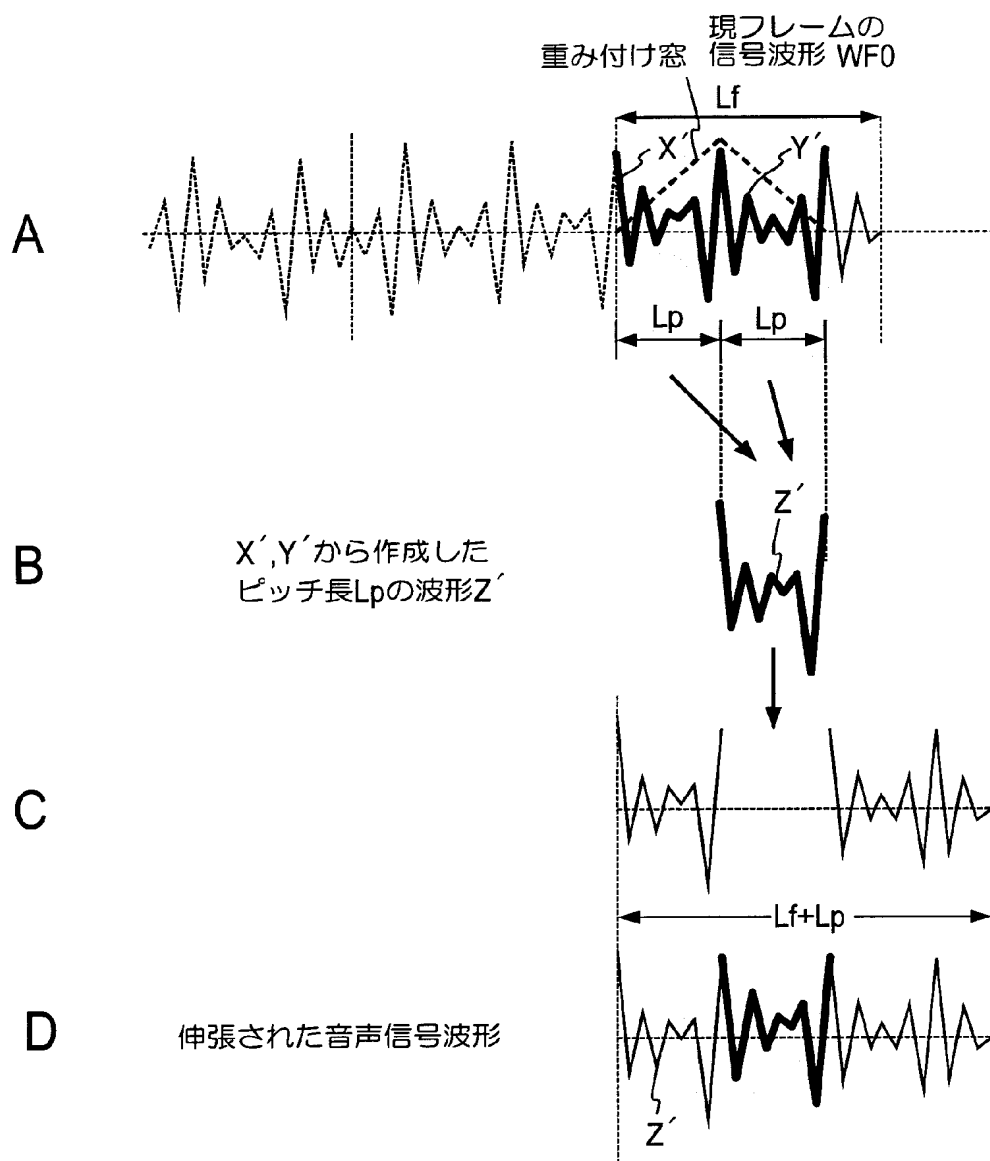
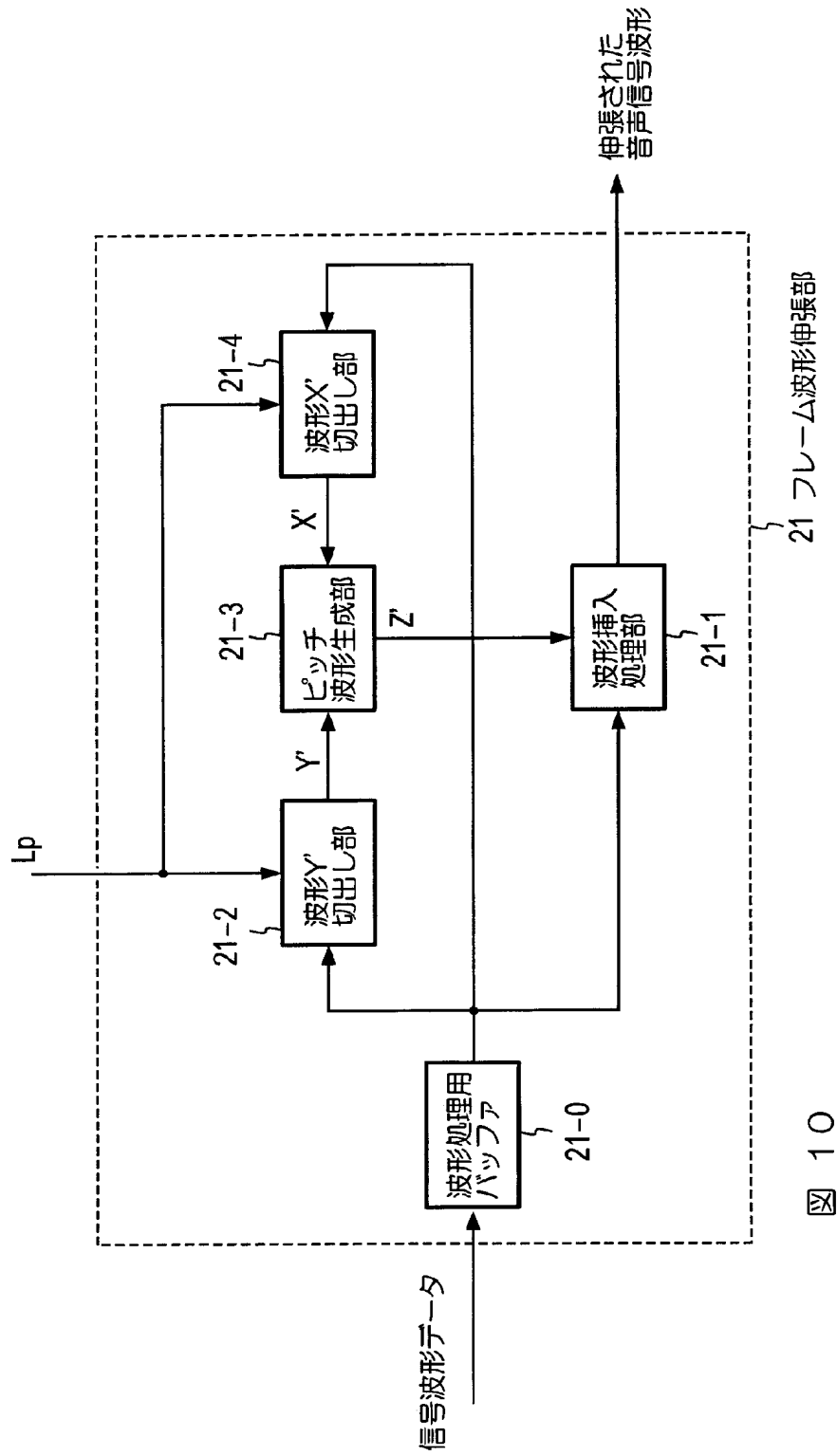
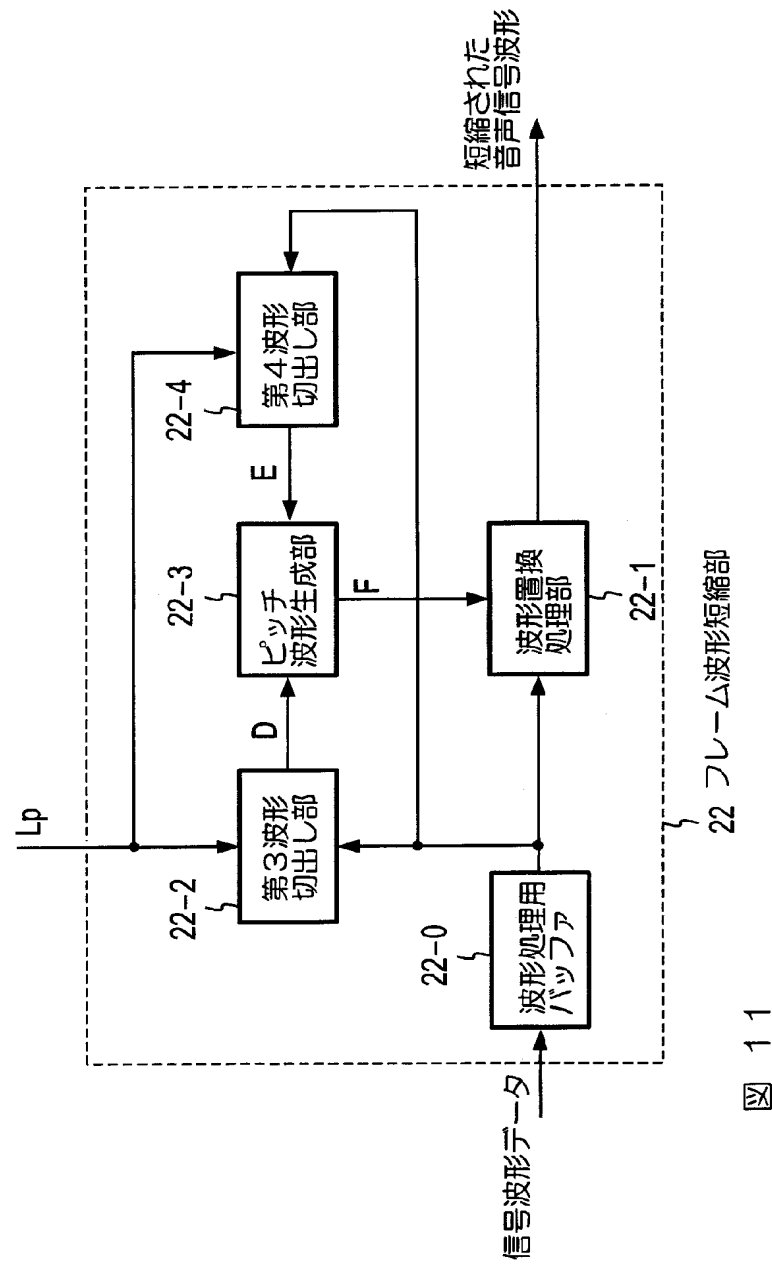


図 9

[図10]



[図11]



[図12]

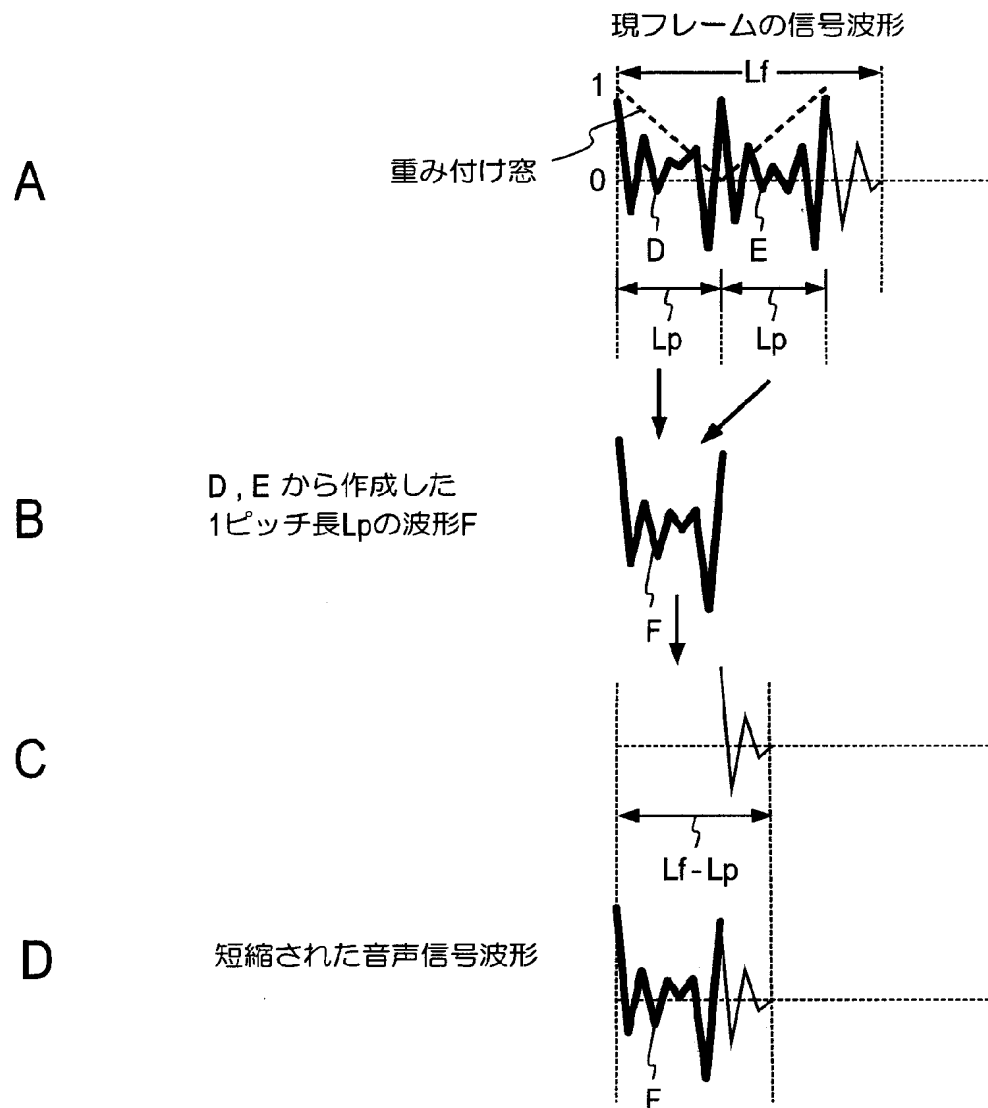


図 12

[図13]

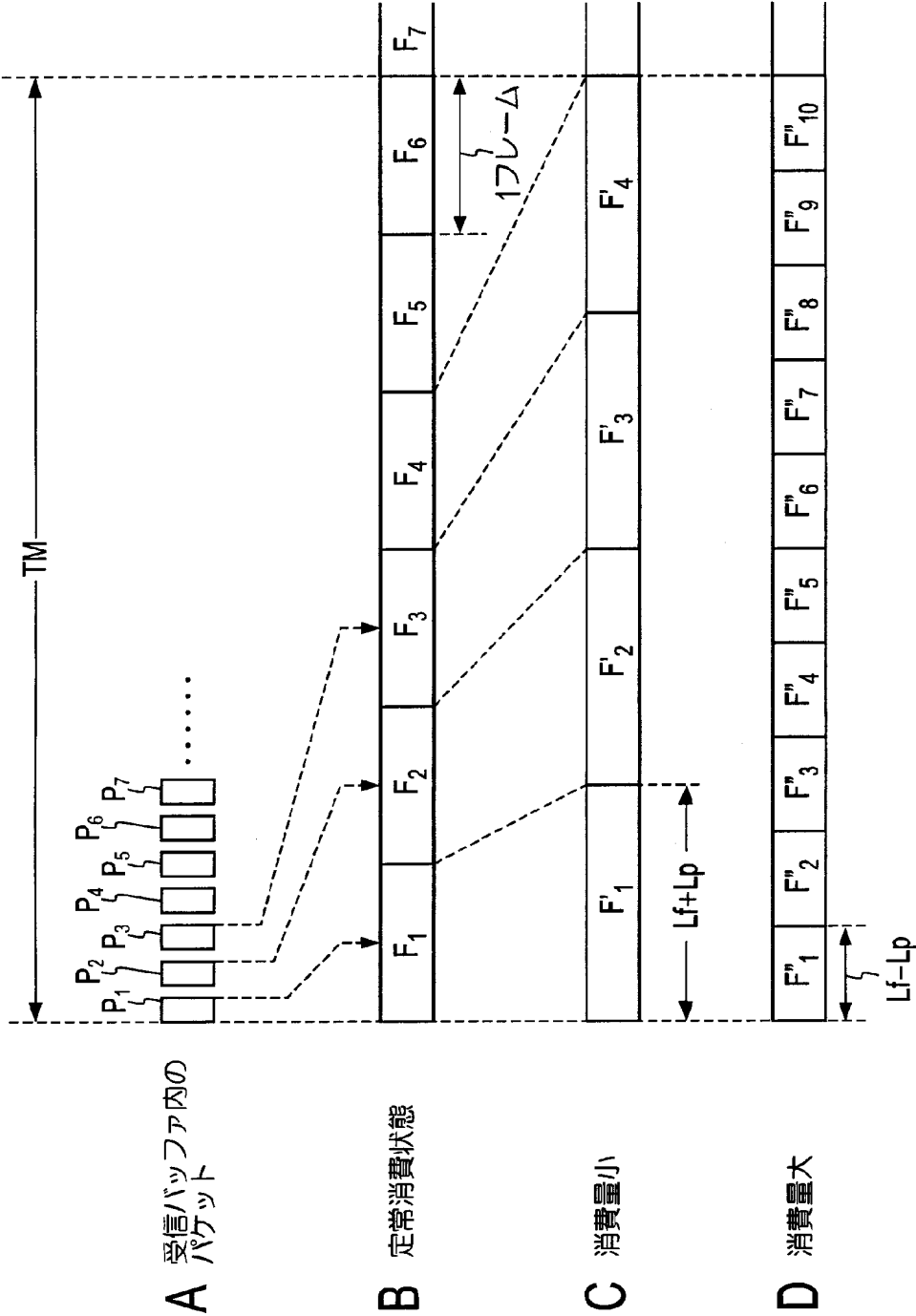


図 13

[図14]

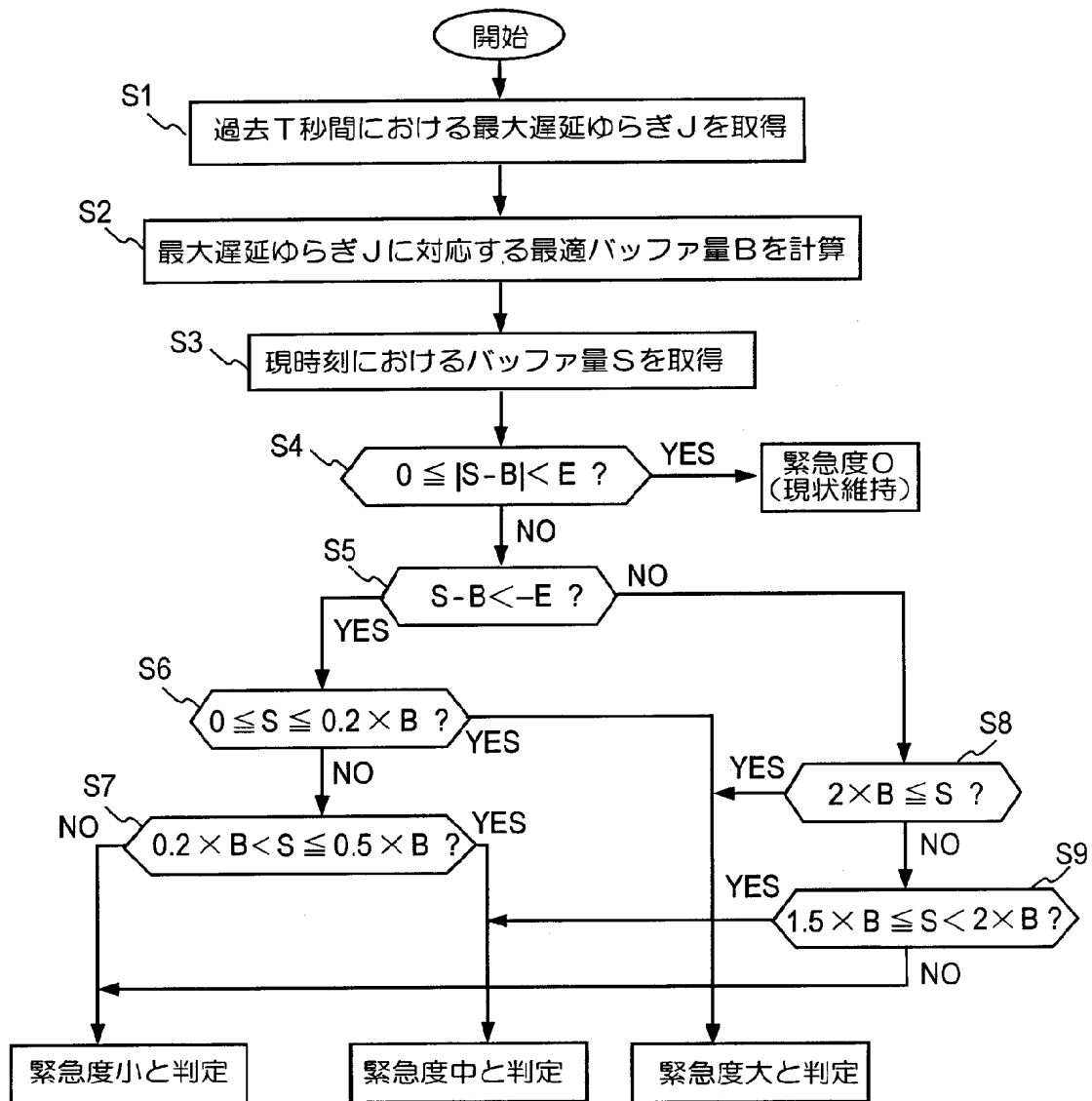


図14

[図15]

表 1

	音声区間	非音声区間
緊急度大	A側またはC側にセット	
緊急度中		
緊急度小	B位置に固定	A側またはC側にセット
緊急度O（現状維持）	B位置に固定	

図 1 5

[図16]

表 2

	音声区間	非音声区間
緊急度大	A側またはC側にセット	
緊急度中		
緊急度小	N1フレームに1回の割合でA側またはC側にセット、それ以外のフレームではB位置に固定（例. N1=5）	A側またはC側にセット
緊急度O（現状維持）	B位置に固定	

図 1 6

[図17]

表 3

	音声区間	非音声区間
緊急度大	A側またはC側にセット	
緊急度中	N2フレームに1回の割合でA側またはC側にセット、それ以外のフレームではB位置に固定（例. N2=2）	A側またはC側にセット
緊急度小	N1フレームに1回の割合でA側またはC側にセット、それ以外のフレームではB位置に固定（例. N1=5）	A側またはC側にセット
緊急度O（現状維持）	B位置に固定	

図 1 7



[図18]

表 4

	音声区間	非音声区間
緊急度大	N3フレームに1回の割合でA側またはC側にセット、それ以外のフレームではB位置に固定（例. N3=1）	
緊急度中	N2フレームに1回の割合でA側またはC側にセット、それ以外のフレームではB位置に固定（例. N2=2）	N4フレームに1回の割合でA側またはC側にセット、それ以外のフレームではB位置に固定（例. N4=1）
緊急度小	N1フレームに1回の割合でA側またはC側にセット、それ以外のフレームではB位置に固定（例. N1=5）	N5フレームに1回の割合でA側またはC側にセット、それ以外のフレームではB位置に固定（例. N5=2）
緊急度0（現状維持）	B位置に固定	

図 18

[図19]

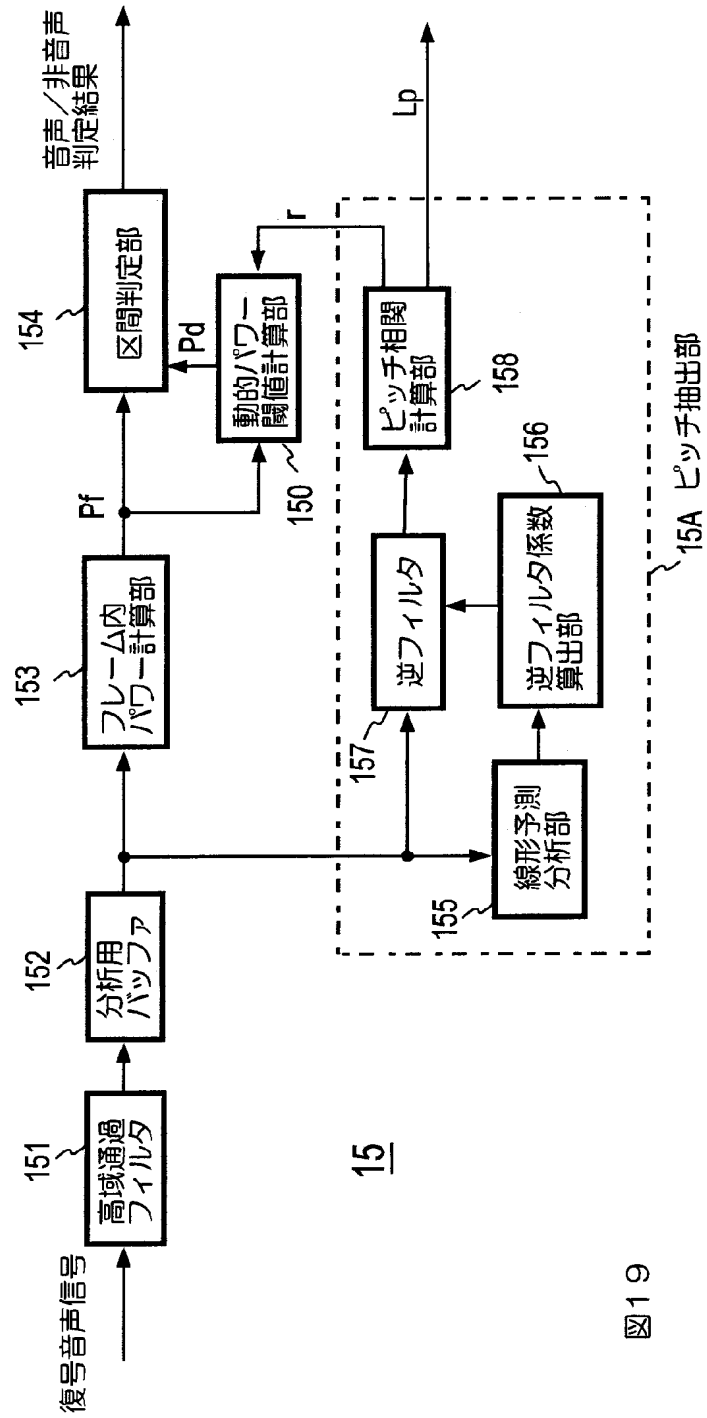


図19

[図20]

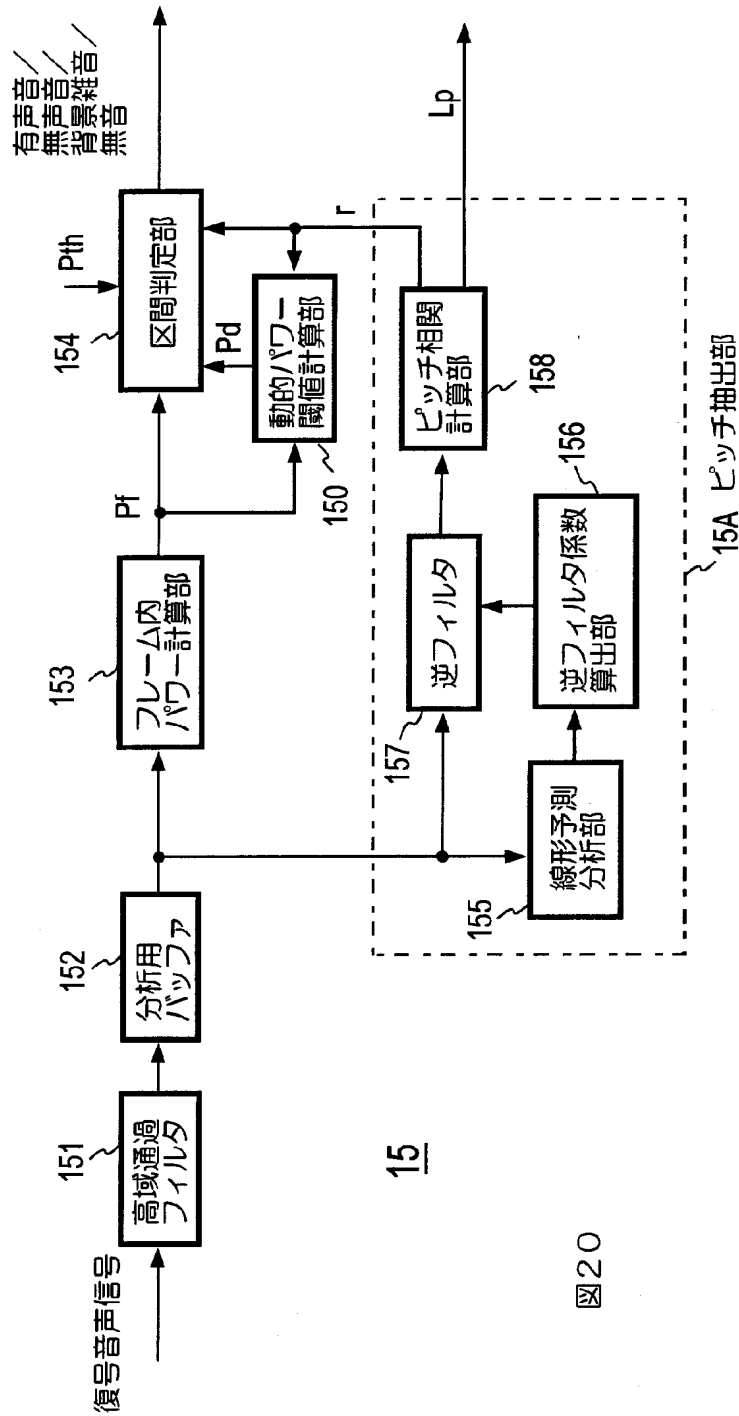


図20

[図21]

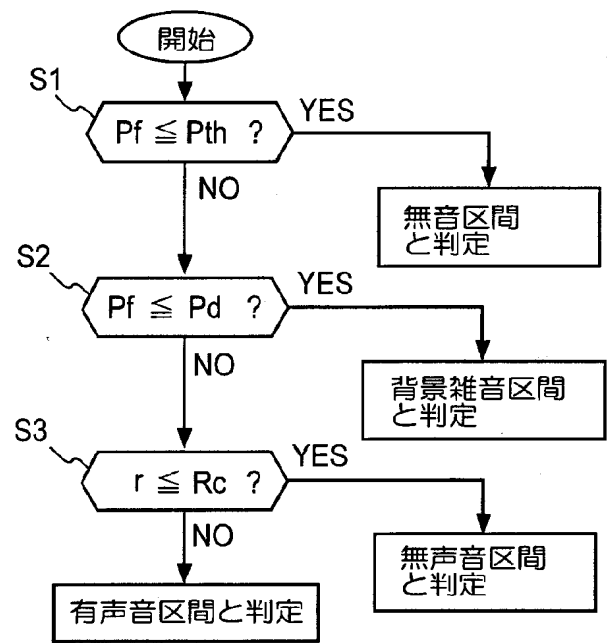


図 2 1

[図22]

表 5

	有声音区間	無声音区間	背景雑音区間	無音区間
緊急度大	A側またはC側にセット			
緊急度中	セット割合N6	セット割合N7	セット割合N8	セット割合N9
緊急度小	セット割合N10	セット割合N11	セット割合N12	セット割合N13
緊急度0	B位置に固定（現状維持）			

図 2 2

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2005/009569

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> H04L12/56, G10L19/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> H04L12/56, G10L19/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2005
Kokai Jitsuyo Shinan Koho	1971-2005	Toroku Jitsuyo Shinan Koho	1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JSTPlus (JOIS), IEEE Xplore

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y A	JP 2002-164921 A (Oki Electric Industry Co., Ltd.), 07 June, 2002 (07.06.02), Full text; Figs. 1 to 9 & US 2002/0064158 A1	1, 7, 13, 14 2, 5, 8, 11 3, 4, 6, 9, 10, 12
Y A	JP 1-93795 A (Nippon Hoso Kyokai), 12 April, 1989 (12.04.89), Full text; Figs. 1 to 5 (Family: none)	2, 5, 8, 11 1, 3, 4, 6, 7, 9, 10, 12-14



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

08 August, 2005 (08.08.05)

Date of mailing of the international search report

23 August, 2005 (23.08.05)

Name and mailing address of the ISA/

Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl.<sup>7</sup> H04L12/56, G10L19/00

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl.<sup>7</sup> H04L12/56, G10L19/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2005年
日本国実用新案登録公報	1996-2005年
日本国登録実用新案公報	1994-2005年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JSTPlus (JOIS)  
IEEE Xplore

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X Y A	J P 2002-164921 A (沖電気工業株式会社) 2002.06.07, 全文, 第1-9図 & US 2002/0064158 A1	1, 7, 13, 14 2, 5, 8, 11 3, 4, 6, 9, 10, 12
Y A	J P 1-93795 A (日本放送協会) 1989.04.12, 全文, 第1-5図 (ファミリーなし)	2, 5, 8, 11 1, 3, 4, 6, 7, 9, 10, 12-14

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの  
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
「O」口頭による開示、使用、展示等に言及する文献  
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
「&」同一パテントファミリー文献

国際調査を完了した日

08.08.2005

国際調査報告の発送日

23.8.2005

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

荏原 雄一

電話番号 03-3581-1101 内線 3541

5Z

3148